





```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-996-617-10
```

```
Query Match          51.4%; Score 322; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 110 CGGCTGCAGCGGGGTGAGCGGGCGGCGGAGATCTTGAGCCATGGGCGGCGC 169
    |||||||
DB 734 GGGCTGCAGCGGGGTGAGCGGGCGGCGGAGATCTTGAGCCATGGGCGGCGC 675
OY 170 GCGAGCCATCTTGATGCGCTGAGAACCTGACCGCGAGAGCTTCAGAAATTCAAC 229
    |||||||
DB 674 GCGAGCCATCTTGATGCGCTGAGAACCTGACCGCGAGAGCTTCAGAAATTCAAC 615
OY 230 TGAAGCTGCTGCGTGGCTGCGGAGGCTAGCGGCGATCCGCGGGGCGGCTGC 289
    |||||||
DB 614 TGAAGCTGCTGCGTGGCTGCGGAGGCTAGCGGCGATCCGCGGGGCGGCTGC 555
OY 290 TGTCCATGAGACCTTGAGACCTCACCAGACCTGTGACCTTCTACCTGAGACCTAG 349
    |||||||
DB 554 TGTCCATGAGACCTTGAGACCTCACCAGACCTGTGACCTTCTACCTGAGACCTAG 495
OY 350 GCGCGGAGCTACCGCTGAGCTGCTGCGGAGATGGGCTGAGAGATGGCGGCGAGC 409
    |||||||
DB 494 GCGCGGAGCTACCGCTGAGCTGCTGCGGAGATGGGCTGAGAGATGGCGGCGAGC 435
OY 410 TGCAGGCGGCGCACGACCAAGG 431
    |||||||
DB 434 TGCAGGCGGCGCACGACCAAGG 413
```

```
RESULT 5
US-09-841-879B-4
; Sequence 4, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; PRIORITY FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-841-879B-4
```

```
Query Match          51.4%; Score 322; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 110 CGGCTGCAGCGGGGTGAGCGGGCGGCGGAGATCTTGAGCCATGGGCGGCGC 169
    |||||||
DB 7 CGGCTGCAGCGGGGTGAGCGGGCGGCGGAGATCTTGAGCCATGGGCGGCGC 66
OY 170 GCGAGCCATCTTGATGCGCTGAGAACCTGACCGCGAGAGCTTCAGAAATTCAAC 229
    |||||||
DB 67 GCGAGCCATCTTGATGCGCTGAGAACCTGACCGCGAGAGCTTCAGAAATTCAAC 126
```

```
OY 230 TGAAGCTGCTGCGTGGCTGCGGAGGCTAGCGGCGATCCCGGGGCGGCGCTGC 289
    |||||||
DB 127 TGAAGCTGCTGCGTGGCTGCGGAGGCTAGCGGCGGCGATCCCGGGGCGGCGCTGC 186
OY 290 TGTCCATGAGACCTTGAGACCTCACCAGACCTGTGACCTTCTACCTGAGACCTAG 349
    |||||||
DB 187 TGTCCATGAGACCTTGAGACCTCACCAGACCTGTGACCTTCTACCTGAGACCTAG 246
OY 350 GCGCGGAGCTCACCGCTTGAAGCTGCGGAGCATATGGGCTGAGAGATGGCGGCGAGC 409
    |||||||
DB 247 GCGCGGAGCTCACCGCTTGAAGCTGCGGAGCATATGGGCTGAGAGATGGCGGCGAGC 306
OY 410 TGCAGGCGGCGCACGACCAAGG 431
    |||||||
DB 307 TGCAGGCGGCGCACGACCAAGG 328
```

```
RESULT 6
US-09-841-879B-19/c
; Sequence 19, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES TH
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; PRIORITY FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879B-19
```

```
Query Match          51.4%; Score 322; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 110 CGGCTGCAGCGGGGTGAGCGGGCGGCGGAGATCTTGAGCCATGGGCGGCGC 169
    |||||||
DB 734 GCGCTGCAGCGGGGTGAGCGGGCGGCGGAGATCTTGAGCCATGGGCGGCGC 675
OY 170 GCGAGCCATCTTGATGCGCTGAGAACCTGACCGCGAGAGCTTCAGAAATTCAAC 229
    |||||||
DB 674 GCGAGCCATCTTGATGCGCTGAGAACCTGACCGCGAGAGCTTCAGAAATTCAAC 615
OY 230 TGAAGCTGCTGCGTGGCTGCGGAGGCTAGCGGCGATCCCGGGGCGGCGCTGC 289
    |||||||
DB 614 TGAAGCTGCTGCGTGGCTGCGGAGGCTAGCGGCGATCCCGGGGCGGCGCTGC 555
OY 290 TGTCCATGAGACCTTGAGACCTCACCAGACCTGTGACCTTCTACCTGAGACCTAG 349
    |||||||
DB 554 TGTCCATGAGACCTTGAGACCTCACCAGACCTGTGACCTTCTACCTGAGACCTAG 495
OY 350 GCGCGGAGCTCACCGCTTGAAGCTGCGGAGCATATGGGCTGAGAGATGGCGGCGAGC 409
    |||||||
DB 494 GCGCGGAGCTCACCGCTTGAAGCTGCGGAGCATATGGGCTGAGAGATGGCGGCGAGC 435
OY 410 TGCAGGCGGCGCACGACCAAGG 431
    |||||||
DB 434 TGCAGGCGGCGCACGACCAAGG 413
```

```
RESULT 7
US-09-925-299-648
; Sequence 648, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
```



```

RESULT 10
US-09-841-879B-6
; Sequence 6, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879B-6

Query Match
Best Local Similarity 100.0%; Pred. No. 9.4e-50;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 ATGGGGCGGGCGGCGGACGCGCATCTGATGGCGGCGTGGAGAACTGACCGCGGAGAGCTC 216
Db 1 ATGGGGCGGGCGGCGGACGCGCATCTGATGGCGGCGTGGAGAACTGACCGCGGAGAGCTC 60

QY 217 AAGAAGTTCAGCTGAAGCTGCTGTCGGTCCGCTGCGCGGAGGGCTACGGGCGCATCCCG 276
Db 61 AAGAAGTTCAGCTGAAGCTGCTGTCGGTCCGCTGCGCGGAGGGCTACGGGCGCATCCCG 120

QY 277 CGGGGGCGGCGTGTCTCCATGAGAGCGCTTGGAGCTCACCAGCAAGCTGGTCTCTAC 336
Db 121 CGGGGGCGGCGTGTCTCCATGAGAGCGCTTGGAGCTCACCAGCAAGCTGGTCTCTAC 180

QY 337 CTGGAGACTCAGCGGGCGGAGCTCACCCTAACGTGCTGCGCGGAGGCTCTGCAGAG 396
Db 181 CTGGAGACTCAGCGGGCGGAGCTCACCCTAACGTGCTGCGCGGAGGCTCTGCAGAG 240

QY 397 ATGGCGGGGCGAGCTGCGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 431
Db 241 ATGGCGGGGCGAGCTGCGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 275

RESULT 11
US-09-728-721-60
; Sequence 60, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (89)...(667)
US-09-728-721-60

Query Match
Best Local Similarity 28.6%; Score 178.8; DB 10; Length 777;
Matches 251; Conservative 0; Mismatches 102; Indels 1; Gaps 1;

QY 81 CGCCCGGAGCCGACCTTCCTCCGTGGCGGCGTGCAGC-GGGGTGAGCGGCGGCGG 139
Db 12 GCTCCGCGAGCAGCAGCGCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 71

QY 140 CCGGGATCTCTGAGCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 199
Db 72 CTGCCACCCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 131

QY 200 TGACCGCGGAGGAGCTCAAGATTCAAGCTGAGCTGCTGCGGCGGCGGCGGCGG 259
Db 132 TGTGAGGGATGAACTCAAAAAGTTCAAGATTCAAGCTGAGCTGAGCTGAGCTGAG 191

QY 260 GCTACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 319
Db 192 GCTATGGCGCATCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 251

QY 320 AGCTGTACGCTTCTACTGAGACTTACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 379
Db 252 AACTGTACGCTTCTACTGAGACTTACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 311

QY 380 ACATGGGCTGCGAGGATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 433
Db 312 ACATGGGCTTACGAGAGCTGCTGAGGAGCTGCAACGACTAAGAAAGAGTCTG 365

RESULT 12
US-09-841-879B-1
; Sequence 1, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES TH
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(667)
US-09-841-879B-1

Query Match
Best Local Similarity 70.9%; Pred. No. 1.7e-29;
Matches 251; Conservative 0; Mismatches 102; Indels 1; Gaps 1;

QY 81 CGCCCGGAGCCGACCTTCCTCCGTGGCGGCGTGCAGC-GGGGTGAGCGGCGGCGG 139
Db 12 GCTCCGCGAGCAGCAGCGCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 71

QY 140 CCGGGATCTCTGAGCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 199
Db 72 CTGCCACCCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 131

QY 200 TGACCGCGGAGGAGCTCAAGATTCAAGCTGAGCTGCTGCGGCGGCGGCGGCGG 259
Db 132 TGTGAGGGATGAACTCAAAAAGTTCAAGATTCAAGCTGAGCTGAGCTGAGCTGAG 191

```









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:21:46 ; Search time 2672 seconds

(without alignments)  
6818.247 Million cell updates/sec

Title: US-09-691-763b-4

Sequence: 1 aggccttctgtggaagc.....ggaaggaaggaagcattcatt 626

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_da: \*  
2: gb\_hg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_pi: \*  
8: gb\_pi: \*  
9: gb\_pi: \*  
10: gb\_pi: \*  
11: gb\_pi: \*  
12: gb\_pi: \*  
13: gb\_pi: \*  
14: gb\_pi: \*  
15: gb\_pi: \*  
16: gb\_pi: \*  
17: gb\_pi: \*  
18: gb\_pi: \*  
19: gb\_pi: \*  
20: gb\_pi: \*  
21: gb\_pi: \*  
22: gb\_pi: \*  
23: gb\_pi: \*  
24: gb\_pi: \*  
25: gb\_pi: \*  
26: gb\_pi: \*  
27: gb\_pi: \*  
28: gb\_pi: \*  
29: gb\_pi: \*  
30: gb\_pi: \*  
31: gb\_pi: \*  
32: gb\_pi: \*  
33: gb\_pi: \*  
34: gb\_pi: \*  
35: gb\_pi: \*  
36: gb\_pi: \*  
37: gb\_pi: \*  
38: gb\_pi: \*  
39: gb\_pi: \*  
40: gb\_pi: \*  
41: gb\_pi: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	626	100.0	626	AX118621	AX118621 Sequence
2	626	100.0	2821	AX118618	AX118618 Sequence
3	626	100.0	2821	AF184072	AF184072 Homo sapi
4	626	100.0	23305	AC009088	AC009088 Homo sapi
5	368	58.8	779	AX017270	AX017270 Sequence
6	361	57.7	782	AX459863	AX459863 Sequence
7	361	57.7	782	AB023416	AB023416 Homo sapi
8	342.8	54.8	713	AX118641	AX118641 Sequence
9	342.8	54.7	405	AF255794	AF255794 Homo sapi
10	342.6	54.7	770	AX118643	AX118643 Sequence
11	342.6	54.7	770	AX118619	AX118619 Sequence
12	342.6	54.7	770	AF184073	AF184073 Homo sapi
13	340	54.3	340	AX118622	AX118622 Sequence
14	322	51.4	740	AX082246	AX082246 Homo sapi
15	322	51.4	740	AF384665	AF384665 Homo sapi
16	301	48.1	740	AK000211	AK000211 Homo sapi
17	275	43.9	585	AX082248	AX082248 Sequence
18	275	43.9	712	AF310103	AF310103 Homo sapi
19	268	42.8	772	BC013569	BC013569 Homo sapi
20	232.4	37.1	768	BC004470	BC004470 Homo sapi
21	179	28.6	1115	AB059327	AB059327 Mus muscu
22	179	28.6	7117	AC101465	AC101465 Mus muscu
23	179	28.6	191928	AC124461	AC124461 Mus muscu
24	179	28.6	225984	AC093175	AC093175 Mus muscu
25	178.8	28.6	777	AX082258	AX082258 Sequence
26	178.6	28.5	732	BC008252	BC008252 Mus muscu
27	178.6	28.5	770	AB032249	AB032249 Mus muscu
28	178.6	28.5	803	AX118637	AX118637 Sequence
29	177.6	28.4	551	AX459859	AX459859 Sequence
30	175.6	28.1	701	AF310104	AF310104 Mus muscu
31	174.4	27.9	556	AF086332	AF086332 Homo sapi
32	174.4	27.9	587	AF454669	AF454669 Homo sapi
33	174.4	27.9	807	AX048091	AX048091 Sequence
34	169.4	27.1	270	AX417218	AX417218 Sequence
35	166	26.5	168975	AC117170	AC117170 Rattus no
36	165	26.4	579	AX082260	AX082260 Sequence
37	129	20.6	52521	AC106629	AC106629 Rattus no
38	117.8	18.8	605	AX118639	AX118639 Sequence
39	86.2	13.8	16891	AF111163	AF111163 Homo sapi
40	86.2	13.8	168700	AC009115	AC009115 Homo sapi
41	86.2	13.8	239566	HSARJ03147	HSARJ03147 Homo sapi
42	74.8	11.9	2825	HSARENO	Y14441 Homo sapien
43	74.8	11.9	3500	AX459861	AX459861 Sequence
44	74.8	11.9	3511	AF018080	AF018080 Homo sapi
45	73	11.7	95038	AF015416	AF015416 Homo sapi

# ALIGNMENTS

RESULT 1  
AX118621  
LOCUS AX118621 626 bp DNA linear PAT 11-MAY-2001  
DEFINITION Sequence 4 from Patent WO0129235.  
ACCESSION AX118621  
VERSION AX118621.1 GI:14035572  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Vertino, P.M.  
TITLE Tm1 compositions and methods of use  
JOURNAL Patent: WO 0129235-A 4 26-APR-2001;

Pred. No. is the number of results predicted by chance to have a

FEATURES Emory University (US)  
SOURCE 1. 626 Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 103 a 216 c 216 g 91 t  
ORIGIN

Query Match 100.0%; Score 626; DB 6; Length 626;  
Best Local Similarity 100.0%; Pred. No. 1.8e-75;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTTTTGTCTGAGGCAACGACCGGGGCGGGAGTCCGGAGACCAAGTGGAGG 60  
DB 1 AGCGCTTTTGTCTGAGGCAACGACCGGGGCGGGAGTCCGGAGACCAAGTGGAGG 60  
QY 61 AAGCGGGGAGTCCAGGTTCCGCCCGGAGCGACTTCTCTGTCGGGCGGCTGACAG 120  
DB 61 AAGCGGGGAGTCCAGGTTCCGCCCGGAGCGACTTCTCTGTCGGGCGGCTGACAG 120  
QY 121 GGGTGAAGCGGCGGAGCGGCGGGATCTTGAGCCATGGGGCGGCGGCGAGCCATC 180  
DB 121 GGGTGAAGCGGCGGAGCGGCGGGATCTTGAGCCATGGGGCGGCGGCGAGCCATC 180  
QY 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAAGTTCAAGCTGAAGCTGCTG 240  
DB 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAAGTTCAAGCTGAAGCTGCTG 240  
QY 241 TCGGTGCGCGCTGGGCGAGGCTACGGGCGCATCCCGGCGGCGGCTGCTGCAATGAC 300  
DB 241 TCGGTGCGCGCTGGGCGAGGCTACGGGCGCATCCCGGCGGCGGCTGCTGCAATGAC 300  
QY 301 GCCTTGAGACCTCAACGCAAGCTGGTACGTTCTACCTGAGAGACTACGCGCGGAGCTC 360  
DB 301 GCCTTGAGACCTCAACGCAAGCTGGTACGTTCTACCTGAGAGACTACGCGCGGAGCTC 360  
QY 361 ACCGCTTAACGTCGTGGCGGACATGGGCTGAGAGAGATGGCCGGAGAGTGCAGGCGGCC 420  
DB 361 ACCGCTTAACGTCGTGGCGGACATGGGCTGAGAGAGATGGCCGGAGAGTGCAGGCGGCC 420  
QY 421 ACGCAGCAGGATAGCGCGCGCGCTTCCCTCCACCGCGTCTTCCCTCCACCGAGCAGC 480  
DB 421 ACGCAGCAGGATAGCGCGCGCGCTTCCCTCCACCGCGTCTTCCCTCCACCGAGCAGC 480  
QY 481 AGCGCTTAACGTCGTGGCGGCTTCCCTGCTTCTCTACCCCTAAACAAAGCTGCTC 540  
DB 481 AGCGCTTAACGTCGTGGCGGCTTCCCTGCTTCTCTACCCCTAAACAAAGCTGCTC 540  
QY 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGAGCAAGGAGACCGCGGCCAGCGC 600  
DB 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGAGCAAGGAGACCGCGGCCAGCGC 600  
QY 601 GGGAGGGAAGGGAAGGAGATCACTT 626  
DB 601 GGGAGGGAAGGGAAGGAGATCACTT 626

RESULT 2  
AX118618 2821 bp DNA linear PAT 11-MAY-2001  
LOCUS AX118618  
DEFINITION Sequence 1 from Patent WO0129235.  
ACCESSION AX118618  
VERSION AX118618.1 GI:14035569  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 2821)  
AUTHORS Vertino,P.M.  
Tms1 compositions and methods of use  
JOURNAL Patent: WO 0129235-A 1 26-APR-2001;  
Emory University (US)

FEATURES Location/Qualifiers  
SOURCE 1. 2821  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 614 a 864 c 757 g 586 t  
ORIGIN

Query Match 100.0%; Score 626; DB 6; Length 2821;  
Best Local Similarity 100.0%; Pred. No. 1.2e-75;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTTTTGTCTGAGGCAACGACCGGGGCGGGAGTCCGGAGACCAAGTGGAGG 60  
DB 1 AGCGCTTTTGTCTGAGGCAACGACCGGGGCGGGAGTCCGGAGACCAAGTGGAGG 1159  
QY 61 AAGCGGGGAGTCCAGGTTCCGCCCGGAGCGACTTCTCTGTCGGGCGGCTGACAG 120  
DB 61 AAGCGGGGAGTCCAGGTTCCGCCCGGAGCGACTTCTCTGTCGGGCGGCTGACAG 1219  
QY 121 GGGTGAAGCGGCGGAGCGGCGGGATCTTGAGCCATGGGGCGGCGGCGAGCCATC 180  
DB 121 GGGTGAAGCGGCGGAGCGGCGGGATCTTGAGCCATGGGGCGGCGGCGAGCCATC 1279  
QY 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAAGTTCAAGCTGAAGCTGCTG 240  
DB 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAAGTTCAAGCTGAAGCTGCTG 1339  
QY 241 TCGGTGCGCGCTGGGCGAGGCTACGGGCGCATCCCGGCGGCGGCTGCTGCAATGAC 300  
DB 241 TCGGTGCGCGCTGGGCGAGGCTACGGGCGCATCCCGGCGGCGGCTGCTGCAATGAC 1399  
QY 301 GCCTTGAGACCTCAACGCAAGCTGGTACGTTCTACCTGAGAGACTACGCGCGGAGCTC 360  
DB 301 GCCTTGAGACCTCAACGCAAGCTGGTACGTTCTACCTGAGAGACTACGCGCGGAGCTC 1459  
QY 361 ACCGCTTAACGTCGTGGCGGACATGGGCTGAGAGAGATGGCCGGAGAGTGCAGGCGGCC 420  
DB 361 ACCGCTTAACGTCGTGGCGGACATGGGCTGAGAGAGATGGCCGGAGAGTGCAGGCGGCC 1519  
QY 421 ACGCAGCAGGATAGCGCGCGCGCTTCCCTCCACCGCGTCTTCCCTCCACCGAGCAGC 480  
DB 421 ACGCAGCAGGATAGCGCGCGCGCTTCCCTCCACCGCGTCTTCCCTCCACCGAGCAGC 1579  
QY 481 AGCGCTTAACGTCGTGGCGGCTTCCCTGCTTCTCTACCCCTAAACAAAGCTGCTC 540  
DB 481 AGCGCTTAACGTCGTGGCGGCTTCCCTGCTTCTCTACCCCTAAACAAAGCTGCTC 1639  
QY 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGAGCAAGGAGACCGCGGCCAGCGC 600  
DB 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGAGCAAGGAGACCGCGGCCAGCGC 1699  
QY 601 GGGAGGGAAGGGAAGGAGATCACTT 626  
DB 601 GGGAGGGAAGGGAAGGAGATCACTT 1725

RESULT 3  
AF184072 2821 bp DNA linear PRI 02-FEB-2001  
LOCUS AF184072  
DEFINITION Homo sapiens target of methylation-induced silencing 1 (TMS1) gene,  
complete cds.  
ACCESSION AF184072  
VERSION AF184072.1 GI:9863861  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 2821)  
AUTHORS Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T.  
and Vertino,P.M.  
TMS1, a novel proapoptotic caspase recruitment domain protein, is a  
target of methylation-induced gene silencing in human breast

JOURNAL Cancer Res. 60 (22), 6236-6242 (2000)  
MEDLINE 20552139  
PUBMED 11103776  
REFERENCE 2 (bases 1 to 2821)  
AUTHORS McConell,B.B. and Vertino,P.M.  
TITLE Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1  
JOURNAL Cancer Res. 60 (22), 6243-6247 (2000)  
MEDLINE 20552140  
PUBMED 11103777  
REFERENCE 3 (bases 1 to 2821)  
AUTHORS Vertino,P.M.  
TITLE Direct Submission  
JOURNAL Submitted (09-SEP-1999) Radiation Oncology, Emory University School of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA

FEATURES  
source 1..2821  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="sperm"  
/clone\_lib="Research Genetics CTRB-HSP-C BAC library"  
complement(44..93)  
/note="LINE"  
/rpl\_family="L2"  
/rpl\_type-dispersed complement(95..256)  
/note="SINE"  
/rpl\_family="Alu"  
/rpl\_type-dispersed complement(257..408)  
/note="LINE"  
/rpl\_family="L2"  
/rpl\_type-dispersed 1111..1722  
/note="CpG island"  
1190..2460  
/gene="TMS1"  
join(1190..1529,1743..1799,2105..2460)  
/gene="TMS1"  
/product="target of methylation-induced silencing 1"  
1190..1529  
/gene="TMS1"  
/number=1  
join(1256..1529,1743..1799,2105..2361)  
/gene="TMS1"  
/function="Involved in apoptosis"  
/note="contains caspase-recruitment domain; CARD protein; target of methylation-mediated gene silencing in Homo sapien tumors"  
/codon\_start=1  
/product="target of methylation-induced silencing 1"  
/protein\_id="AAC01187.1"  
/db\_xref="GI:9863862"  
/translation="MGRADATLDALENTAEEIKKFKLKLSVPLRESYGRTRPGAL  
LSMDALIDIKIVSEFLRYGAEITANVLRDMGLOEAGQLAAHQSGAAPAGTQA  
PPOSAKPGIHLIDQIRALIRAVNEMILDALYGVLTDEQYQAVRAEPNSKMR  
KLFSFTPANMTCKDILLOALRESQSYVEDLERS"  
1743..1799  
/gene="TMS1"  
/number=2  
2105..2460  
/gene="TMS1"  
/number=3  
2438..2443  
/gene="TMS1"  
polya\_signal  
/gene="TMS1"  
BASE COUNT 614 a 864 c 757 g 586 t  
ORIGIN

Query Match 100.0%; Score 626; DB 9; Length 2821;  
Best Local Similarity 100.0%; Pred. No. 1.2e-75;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTTTTCTGTCAGAGCAACGACCGGCGGAGTTCGGAGACAGTGGAGG 60  
|||||  
Db 1100 AGCGCTTTTCTGTCAGAGCAACGACCGGCGGAGTTCGGAGACAGTGGAGG 1159  
QY 61 AAGCGGGAGACTCCAGTTCGCCCCCGAGCCGACCTTCTCTGTCGGCGGCGAGCG 120  
|||||  
Db 1160 AAGCGGGAGACTCCAGTTCGCCCCCGAGCCGACCTTCTCTGTCGGCGGCGAGCG 1219  
QY 121 GGGTACCGCGCGACGCGCGCGGAGATCTGAGCCATGGGCGCGCGCGAGCCATC 180  
|||||  
Db 1220 GGGTACCGCGCGACGCGCGCGGAGATCTGAGCCATGGGCGCGCGAGCCATC 1279  
QY 181 CTGATGCGCTGGAAGACTGACCGCCGAGAGTCAAGATTAAGTGAAGTCTG 240  
|||||  
Db 1280 CTGATGCGCTGGAAGACTGACCGCCGAGAGTCAAGATTAAGTGAAGTCTG 1339  
QY 241 TCGTGGCGGTGGCGGAGGCGTACGGGCGCATCCCGGCGCGCGCTGCTCATGAG 300  
|||||  
Db 1340 TCGTGGCGGTGGCGGAGGCGTACGGGCGCATCCCGGCGCGCGCTGCTCATGAG 1399  
QY 301 GCGTTGACCTTCACGACCAAGCTGTCAGCTTCTACTGTGAGACCTACGGGCGGAGCTG 360  
|||||  
Db 1400 GCGTTGACCTTCACGACCAAGCTGTCAGCTTCTACTGTGAGACCTACGGGCGGAGCTG 1459  
QY 361 ACCGCTAACGCTGCTGCGACATGGGCGCTGCGAGATGGCGCGAGCTGACGCGGCC 420  
|||||  
Db 1460 ACCGCTAACGCTGCTGCGACATGGGCGCTGCGAGATGGCGCGAGCTGACGCGGCC 1519  
QY 421 ACGGACCAAGGTCAGGCGCGCGCGCTTCCGCTCCACCCGCTTCCGCTCCACGACG 480  
|||||  
Db 1520 ACGGACCAAGGTCAGGCGCGCGCGCTTCCGCTCCACCCGCTTCCGCTCCACGACG 1579  
QY 481 AGCGTTACCCCGCGGCGCTTCCGCTTCTGTTCTCTACCTTAAACAAAGTGTGTC 540  
|||||  
Db 1580 AGCGTTACCCCGCGGCGCTTCCGCTTCTGTTCTCTACCTTAAACAAAGTGTGTC 1639  
QY 541 TACGGGAAGAGGAGCGCGCGCGCTTACGACGACCAAGGAGACCGCGCGCGACGCG 600  
|||||  
Db 1640 TACGGGAAGAGGAGCGCGCGCGCTTACGACGACCAAGGAGACCGCGCGCGACGCG 1699  
QY 601 GCGAAGGAGGAGGAGGAGGATCATTT 626  
|||||  
Db 1700 GCGAAGGAGGAGGAGGAGGATCATTT 1725

RESULT 4  
AC009088/c 233305 bp DNA linear PRI 03-APR-2002  
LOCUS Homo sapiens chromosome 16 clone RP11-388M20, complete sequence.  
DEFINITION AC009088  
ACCESSION AC009088  
VERSION AC009088.7 GI:1909387  
KEYWORDS HNG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 233305)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 233305)  
TITLE DOE Joint Genome Institute.  
JOURNAL Submitted (03-APR-1999) Production Sequencing Facility, DOE Joint  
AUTHORS Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
TITLE 3 (bases 1 to 233305)  
JOURNAL Direct Submission  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-APR-2002) Production Sequencing Facility, DOE Joint  
AUTHORS Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
TITLE On Apr 3, 2002 this sequence version replaced gi:16924080.  
COMMENT Sequence Quality Assessment:  
This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
Genbank flat file format but are available as part  
of this entry's ASN.1 file.

## FEATURES

Location/Qualifiers  
1. 233305  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-388M20"

## misc\_feature

1. 233305

/note="clone overlap not submitted"

BASE COUNT 54492 a 59306 c 61380 g 58127 t

## ORIGIN

Query Match 100.0%; Score 626; DB 9; Length 233305;  
Best Local Similarity 100.0%; Pred. No. 4e-76;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCGCTTTGCTGAGGAGCAACGACCGGGGCGGAGTGGAGACCAAGTGGAGG 60  
Db 164569 AGCGCCCTTTGCTGAGGAGCAACGACCGGGGCGGAGTGGAGACCAAGTGGAGG 164510  
QY 61 AAGCGGGGAGATCCAGATCTCCCGCCGAGCCGACTCTCTCGTGGCGGGCTGCAAGC 120  
Db 164509 AAGCGGGGAGATCCAGATCTCCCGCCGAGCCGACTCTCTCGTGGCGGGCTGCAAGC 164450  
QY 121 GGGTGAGCGGCGGAGCGGCGGGGATCTGAGACCATGGGCGCGCGGAGCGCATC 180  
Db 164449 GGGTGAGCGGCGGAGCGGCGGGGATCTGAGACCATGGGCGCGCGGAGCGCATC 164390  
QY 181 CTGGATGCGCTGGAACACTGACCGCCGAGAGACTTAAACAATTAACTCAACTCTCTG 240  
Db 164389 CTGGATGCGCTGGAACACTGACCGCCGAGAGACTTAAACAATTAACTCAACTCTCTG 164330  
QY 241 TCGGTCCGCTGCGGAGGAGCTACGGGCGCATCCCGGGGCGCGCTGCTGCATGAGC 300  
Db 164329 TCGGTCCGCTGCGGAGGAGCTACGGGCGCATCCCGGGGCGCGCTGCTGCATGAGC 164270  
QY 301 GCGTTGAGACTACCGCAAACTGCTGACTTCTACTCTGGAACCTTAAGCGCGGAGCTC 360  
Db 164269 GCGTTGAGACTACCGCAAACTGCTGACTTCTACTCTGGAACCTTAAGCGCGGAGCTC 164210  
QY 361 ACCGCTAACGTGCTGCGGACATGGGCTGACAGAGATGGCGGCGAGCTGCAGGCGGC 420  
Db 164209 ACCGCTAACGTGCTGCGGACATGGGCTGACAGAGATGGCGGCGAGCTGCAGGCGGC 164150  
QY 421 ACGGACCAAGGTGACCGCGCCGCTTCCCTCCACCGCGTCTTTCCCTCCACCAACG 480  
Db 164149 ACGGACCAAGGTGACCGCGCCGCTTCCCTCCACCGCGTCTTTCCCTCCACCAACG 164090  
QY 481 ACGGCTTACCCCGCGGCGCTTTCGCTTCTGTTCTCTTACCCCTAAACAACACTGCTC 540  
Db 164089 ACGGCTTACCCCGCGGCGCTTTCGCTTCTGTTCTCTTACCCCTAAACAACACTGCTC 164030  
QY 541 TACCGGAAAGAGAGGTCCCGACGCTTGGCTTACCGCAAGCGGAGCCCGGCGGCGG 600  
Db 164029 TACCGGAAAGAGAGGTCCCGACGCTTGGCTTACCGCAAGCGGAGCCCGGCGGCGG 163970  
QY 601 GGGAGGAGGAGGAGGATCACTT 626  
Db 163969 GGGAGGAGGAGGAGGATCACTT 163944

RESULT 5  
LOCUS AX017270 779 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 21 from Patent WO9947669.  
ACCESSION AX017270

## VERSION

AX017270.1 GI:10042188

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

```
/db_xref="taxon:9606"  
/chromosome="16"
```

```
/note="unnamed protein product
/codon_start=1
```

FEATURES	source
PROTEIN	/protein_id="CAC38476.1" /db_xref="GI:14035593" /translation="MGARADALLENLTAELEKKFKILSVLPREGGRIPIGAL LSNADLIDTLKLYSEETTYCAELTANVLRDGLQDPAQGLQAAHTGHPIDHRAA LIRATVNVREHLDALIKGVYLVLEDTQAVAREFTNPSKKRLFSTFAMNMTCIDLLQ ALRFSQSYLVLEDLRS"
BASE COUNT	148 a 224 c 223 g 118 t
ORIGIN	
Query Match	54.8%; Score 342.8; DB 6; Length 713;
Best Local Similarity	98.0%; Pred. No. 2,5e-37;
Matches 347; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
QY	83 CCCCAGGCGCACTTCCMCGTGTGGGGGGCGGACGAGGGGGGTAGCGCGGACAGGGCGG 142
DB	1 CCACGGCTCCGACTCTCTCTGTGTCGGCGGTGCACGCGGGGTAGCGCGGACGGCGG 60
QY	143 GGGATCTTGAGCCATGGGGCGCGCGCGACGCCATCTGTGATGCGTGAAGACTGA 202
DB	61 GGGATCTTGAGCCATGGGGCGCGCGCGACGCCATCTGTGATGCGTGAAGACTGA 120
QY	203 CCCCAGAGAGCTCAAGAACTTAAAGCTGAAGCTGCTGTGGTGGCCGCGCGAGAGGCT 262
DB	121 CCGCGAGAGAGCTCAAGAACTGAAGCTGAAGCTGTGCTGTGGTGGCCGCGCGAGAGGCT 180
QY	263 ACGGGGCAATCCCGGGGGGGCGGCTGTGCTCATGAGAGCGCTTGGAGCTCAGCGACAAAG 322
DB	181 ACGGGGCAATCCCGGGGGGGCGGCTGTGCTCATGAGAGCGCTTGGAGCTCAGCGACAAAG 240
QY	323 TGGTCAAGCTTACCTGAGACCTACGAGGCGCGAGCTCACCGGCTAAAGCTGTGCGCGACA 382
DB	241 TGGTCAAGCTTACCTGAGACCTACGAGGCGCGAGCTCACCGGCTAAAGCTGTGCGCGACA 300
QY	383 TGGGCTGAGAGAGATGGCGGCGACCTCAAGGGGGGCGACGACGAGGGTGGAGC 436
DB	301 TGGGCTGAGAGAGATGGCGGCGACCTCAAGGGGGGCGACGACGAGGGTGGAGC 354
RESULT 9	
AF25794	713 bp mRNA linear PRI 02-FEB-2001
LOCUS	AF25794
DEFINITION	Human spleen target of methylation-induced silencing-1 (TMS1) mRNA.
ACCESSION	AF25794
VERSION	AF25794.1
KEYWORDS	GI:9802275
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	1 (bases 1 to 713) Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T. and Vertino,P.M.
TITLE	TMS1, a novel proapoptotic caspase recruitment domain protein, is a target of methylation-induced gene silencing in human breast cancers
JOURNAL	Cancer Res. 60 (22), 6236-6242 (2000)
MEDLINE	20552139
PUBMED	1103776
REFERENCE	2 (bases 1 to 713)
AUTHORS	McConnell,B.B. and Vertino,P.M.
TITLE	Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1
JOURNAL	Cancer Res. 60 (22), 6243-6247 (2000)
MEDLINE	20552140
PUBMED	1103777
REFERENCE	3 (bases 1 to 713)
AUTHORS	Vertino,P.M.
TITLE	Direct Submission
JOURNAL	Submitted (13-APR-2000) Radiation Oncology, Emory University School of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA
FEATURES	Location/Qualifiers 1..713

FEATURES	LOCATION/QUALIFIERS
ORGANISM	human.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	Vertlno.P.M.
JOURNAL	1 (bases 1 to 405)
REFERENCE	Tmsi Compositions and methods of use
PATENT	Patent: WO 0129235-A 26 26-APR-2001;
Emory University (US)	
LOCATION/QUALIFIERS	
BASE COUNT	148 a 224 c 223 g 118 t
QUERY MATCH	54.8%; Score 342.8; DB 9; Length 713;
Best Local Similarity	98.0%; Pred. No. 2.5e-37;
Matches 347; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
QY	B3 CCCCAGAGCCACATCTCTCTCTGTCGCGGCGTCGACGCGGAGCGAGCGCGC 142
Db	1 CCACGGCTCCGACTTCTCTCTGTCGCGGCGTCGACGCGGAGCGAGCGCGC 60
QY	143 GGGATCTCGAGCCATGGGCGCGCGCGCGCGACGCGCATCTGATGCGCTGAGAACCTGA 202
Db	61 GGGATCTCGAGCCATGGGCGCGCGCGCGCGACGCGCATCTGATGCGCTGAGAACCTGA 120
QY	203 CCGCGAGAGCCTCAGAACGTTGAACTGCTGTGTCGCGCGCGAGGCGCT 262
Db	121 CCGCGAGAGCCTCAGAACGTTGAACTGCTGTGTCGCGCGCGAGGCGCT 180
QY	263 ACGGCGGATCCCGCGGCGCGCTGCTGTCATGACGCGCTTGAGCTCACGAGCAAGC 322
Db	181 ACGGCGGATCCCGCGGCGCGCTGCTGTCATGACGCGCTTGAGCTCACGAGCAAGC 240
QY	323 TGGTACGCTTACTTACCTGAGACCTCAGCGCGCGCGCGCGCGCTTAAGCTGTCGGGACA 382
Db	241 TGGTACGCTTACTTACCTGAGACCTCAGCGCGCGCGCGCGCGCTTAAGCTGTCGGGACA 300
QY	383 TGGGCTTCAGAGAGATGGCGGCGCGCTGACAGGCGCGCGCGCGAGGCTGAGC 436
Db	301 TGGGCTTCAGAGAGATGGCGGCGCGCTGACAGGCGCGCGCGCGAGGCTGAGC 354
RESULT 10	
LOCUS	AX118643
DEFINITION	Sequence 26 from Patent W00129235.
ACCESSION	AX118643
VERSION	AX118643.1 GI:14035594
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	Vertlno.P.M.
JOURNAL	1 (bases 1 to 405)
REFERENCE	Tmsi Compositions and methods of use
PATENT	Patent: WO 0129235-A 26 26-APR-2001;
Emory University (US)	
LOCATION/QUALIFIERS	

source 1..405  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
CDS  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC38477.1"  
/db\_xref="GI:14035595"  
/translation="MGRRADALIDALENTAELKKFKILISVPLRFGYGRIRGAL  
LSMDALDIDKLVSTFLETYTGAGELTANVLRMDQLQMAQLODAATHQSGAAPPAGIA  
PPOSAAKP"  
BASE COUNT 63 a 139 c 147 g 56 t  
ORIGIN

Query Match 54.7%; Score 342.6; DB 6; Length 405;  
Best Local Similarity 98.9%; Pred. No. 3.1e-37;  
Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 83 CCCCAGAGCCGACTTCCTCTGTTGCGGCGCTGACGCGGGGTGAGCGCGGCGGCGCG 142  
|||  
Db 1 CCACGGCTCCGACTTCCTCTGTTGCGGCGCTGACGCGGGGTGAGCGCGGCGGCGGCGG 60  
|||  
Qy 143 GGGATCCTGGAGCCATGGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 202  
|||  
Db 61 GGGATCCTGGAGCCATGGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 120  
|||  
Qy 203 CCGCCGAGAGCTCAAGAGTTCAAGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCT 262  
|||  
Db 121 CCGCCGAGAGCTCAAGAGTTCAAGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCT 180  
|||  
Qy 263 ACGGGCGCATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 322  
|||  
Db 181 ACGGGCGCATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 240  
|||  
Qy 323 TGCTCAGCTTCTACTGAGACCTACGCGCGGAGCTCACCGCTTACGCTGCGCGACCA 382  
|||  
Db 241 TGCTCAGCTTCTACTGAGACCTACGCGCGGAGCTCACCGCTTACGCTGCGCGACCA 300  
|||  
Qy 383 TGGGCTTGAGAGATGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 431  
|||  
Db 301 TGGGCTTGAGAGATGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 349  
|||

RESULT 11  
AX118619 770 bp DNA linear PAT 11-MAY-2001  
LOCUS  
DEFINITION  
Sequence 2 from Patent WO0129235.  
ACCESSION  
AX118619  
VERSION  
AX118619.1 GI:14035570  
KEYWORDS  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 770)  
AUTHORS  
Vertino,P.M.  
TITLE  
Tms1 compositions and methods of use  
JOURNAL  
Patent: WO 0129235-A 2 26-APR-2001;  
Emory University (US)  
FEATURES  
source  
1..770  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
75..662  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC38477.1"  
/db\_xref="GI:14035571"  
/translation="MGRRADALIDALENTAELKKFKILISVPLRFGYGRIRGAL  
LSMDALDIDKLVSTFLETYTGAGELTANVLRMDQLQMAQLODAATHQSGAAPPAGIA  
PPOSAAKPGLHFDQHPAALIAKRVNEMLRMDALYGYLDEQYQAVRAEPTNPSKMR  
KLFTTPAMWMTCKDILLQALRESQSYLVEDLERS"  
BASE COUNT 157 a 247 c 241 g 125 t

ORIGIN

Query Match 54.7%; Score 342.6; DB 6; Length 770;  
Best Local Similarity 98.9%; Pred. No. 2.0e-37;  
Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 83 CCCCAGAGCCGACTTCCTCTGTTGCGGCGCTGACGCGGGGTGAGCGCGGCGGCGGCGGCG 142  
|||  
Db 1 CCACGGCTCCGACTTCCTCTGTTGCGGCGCTGACGCGGGGTGAGCGCGGCGGCGGCGGCGG 60  
|||  
Qy 143 GGGATCCTGGAGCCATGGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 202  
|||  
Db 61 GGGATCCTGGAGCCATGGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
|||  
Qy 203 CCGCCGAGAGCTCAAGAGTTCAAGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 262  
|||  
Db 121 CCGCCGAGAGCTCAAGAGTTCAAGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 180  
|||  
Qy 263 ACGGGCGCATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 322  
|||  
Db 181 ACGGGCGCATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 240  
|||  
Qy 323 TGCTCAGCTTCTACTGAGACCTACGCGCGGAGCTCACCGCTTACGCTGCGCGACCA 382  
|||  
Db 241 TGCTCAGCTTCTACTGAGACCTACGCGCGGAGCTCACCGCTTACGCTGCGCGACCA 300  
|||  
Qy 383 TGGGCTTGAGAGATGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 431  
|||  
Db 301 TGGGCTTGAGAGATGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 349  
|||

RESULT 12  
AF184073 770 bp mRNA linear PRI 02-FEB-2001  
LOCUS  
DEFINITION  
Homo sapiens target of methylation-induced silencing 1 (TMS1) mRNA,  
complete cds.  
ACCESSION  
AF184073  
VERSION  
AF184073.1 GI:9863863  
KEYWORDS  
SOURCE  
Homo sapiens.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 770)  
AUTHORS  
Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T.  
and Vertino,P.M.  
TITLE  
TMS1, a novel proapoptotic caspase recruitment domain protein, is a  
target of methylation-induced gene silencing in human breast  
cancers  
JOURNAL  
Cancer Res. 60 (22), 6236-6242 (2000)  
MEDLINE  
20552139  
PUBMED  
11103776  
REFERENCE  
2 (bases 1 to 770)  
AUTHORS  
McConnell,B.B. and Vertino,P.M.  
TITLE  
Activation of a caspase-9-mediated apoptotic pathway by subcellular  
redistribution of the novel caspase recruitment domain protein TMS1  
JOURNAL  
Cancer Res. 60 (22), 6243-6247 (2000)  
MEDLINE  
20552140  
PUBMED  
11103777  
REFERENCE  
3 (bases 1 to 770)  
AUTHORS  
Vertino,P.M.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (09-SEP-1999) Radiation Oncology, Emory University School  
of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30355, USA  
FEATURES  
source  
1..770  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:159538; IMAGE:1990931"  
/clone\_1lb="Soares breast 3nHBst; NCL\_cgAP\_Ut3 uterine  
tumor"  
1..770  
/gene="TMS1"

CDS

75..662  
/gene="TMS1"  
/function="Involved in apoptosis"  
/note="contains caspase-recruitment domain; CARD protein;  
target of methylation-mediated gene silencing in Homo  
sapien tumors"  
/codon\_start=1  
/product="target of methylation-induced silencing 1"  
/protein\_id="AF01188.1"  
/db\_xref="GI:9863864"  
/translation="MGRARDIIDLAEENLTAELKFKLLSYPLRGYRIPRGAL  
LSMDALDTDLKLVSYLETYGAEITANVLRDMGLQEMAGQLQAAIHQSGAAPGIA  
PPOSAAKPGELHIDQRAALIARVNEMWLDALYGVLTDRQYAVABEPTNSKMR  
KLFSFTPANMWTCKDLLQALRESQSYLVEIDERS"

BASE COUNT 157 a 247 c 241 g 125 t

ORIGIN

Query Match 54.7%; Score 342.6; DB 9; Length 770;  
Best Local Similarity 98.9%; Pred. No. 2.6e-37;  
Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 83 CCGCGAGCCGACTTCTCTCTGTCGGCGGCTGCAGCGGGGTGAGCGCGCAGCGCG 142  
||| |  
Db 1 CCACGCGTCCACTTCTCTCTGTCGGCGGCTGCAGCGGGGTGAGCGCGCAGCGCG 60  
||| |  
QY 143 GGGATCCGAGGCGATGGGGGGGGGGGAGCGGCATCTTGATGCGCTGGAGAACCTGA 202  
||| |  
Db 61 GGGATCCGAGGCGATGGGGGGGGGGGAGCGGCATCTTGATGCGCTGGAGAACCTGA 120  
||| |  
QY 203 CCGCGAGGAGCTTCAGACCTGCAAGCTGCTGTCGGTCCGCTGCGGAGGGCT 262  
||| |  
Db 121 CCGCGAGGAGCTTCAGACCTGCAAGCTGCTGTCGGTCCGCTGCGGAGGGCT 180  
||| |  
QY 263 ACGGCGCATCCCGGGGGGGGGGGGGCTGTCATGAGGAGCGCTTGAGACCTCAGCGAACGC 322  
||| |  
Db 181 ACGGCGCATCCCGGGGGGGGGGGGGCTGTCATGAGGAGCGCTTGAGACCTCAGCGAACGC 240  
||| |  
QY 323 TGGTCAGCTTCTACCTGAGACCTTACGGCGCGGAGCTTACCGCTTAACGTGTCGGCGAGA 382  
||| |  
Db 241 TGGTCAGCTTCTACCTGAGACCTTACGGCGCGGAGCTTACCGCTTAACGTGTCGGCGAGA 300  
||| |  
QY 383 TGGGCGTTCAGAGAGATGGCGGGGAGCTGTCAGAGCGGCGCAGCAGCCAGGG 431  
||| |  
Db 301 TGGGCGTTCAGAGAGATGGCGGGGAGCTGTCAGAGCGGCGCAGCAGCCAGGG 349  
||| |

RESULT 13  
AX118622 340 bp DNA linear PAT 11-MAY-2001  
LOCUS AX118622  
DEFINITION Sequence 5 from Patent WO0129235.  
ACCESSION AX118622  
VERSION AX118622.1 GI:14035573  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 340)  
AUTHORS Vertino,P.M.  
TITLE Tmsl compositions and methods of use  
JOURNAL Patent: WO 0129235-A 5 26-Apr-2001;  
Emory University (US)  
FEATURES  
source  
1..340  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
67..>339  
/note="unnamed protein product"

CDS  
/codon\_start=1  
/protein\_id="CAC38472.1"  
/db\_xref="GI:14035574"  
/translation="MGRARDIIDLAEENLTAELKFKLLSYPLRGYRIPRGAL  
LSMDALDTDLKLVSYLETYGAEITANVLRDMGLQEMAGQLQAAIHQSGAAPGIA  
PPOSAAKPGELHIDQRAALIARVNEMWLDALYGVLTDRQYAVABEPTNSKMR  
KLFSFTPANMWTCKDLLQALRESQSYLVEIDERS"

BASE COUNT 53 a 112 c 127 g 48 t

ORIGIN

Query Match 54.3%; Score 340; DB 6; Length 340;  
Best Local Similarity 100.0%; Pred. No. 7.2e-37;  
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CCGACTCCCTCCCTGCTGCGGCGGCTGCAGCGGGGTGAGCGCGCGCGGAGATCT 130  
||| |  
Db 1 CCGACTCCCTCCCTGCTGCGGCGGCTGCAGCGGGGTGAGCGCGCGCGGAGATCT 60  
||| |  
QY 151 GGAGCCATGGGGGGCGCGCGCGAGCCATCTGGATGCCCTGAGAACCTGACCGCCAG 210  
||| |  
Db 61 GGAGCCATGGGGGGCGCGCGCGAGCCATCTGGATGCCCTGAGAACCTGACCGCCAG 120  
||| |  
QY 211 GAGCTCAAGAGTTCAGAGCTGAGCTGCTGTCGGTCCCTGCGGAGGCTACGGGCGC 270  
||| |  
Db 121 GAGCTCAAGAGTTCAGAGCTGAGCTGCTGTCGGTCCCTGCGGAGGCTACGGGCGC 180  
||| |  
QY 271 ATCCCGGGGGCGCGCTGCTGTCATGAGAGCCCTTGAGCTCAGCCAGACCTGTCAGC 330  
||| |  
Db 181 ATCCCGGGGGCGCGCTGCTGTCATGAGAGCCCTTGAGCTCAGCCAGACCTGTCAGC 240  
||| |  
QY 331 TTCTACTGAGAGCTTACGGCGCGGAGCTTACCGCTTACGCTGCGGAGATGGGCTG 390  
||| |  
Db 241 TTCTACTGAGAGCTTACGGCGCGGAGCTTACCGCTTACGCTGCGGAGATGGGCTG 300  
||| |  
QY 391 CAGAGATGGCGGGCGAGCTGCAAGCGCGCCAGCAGCAGG 430  
||| |  
Db 301 CAGAGATGGCGGGCGAGCTGCAAGCGCGCCAGCAGCAGG 340  
||| |

RESULT 14  
AX082246 740 bp DNA linear PAT 27-FEB-2001  
LOCUS AX082246  
DEFINITION Sequence 48 from Patent WO0100826.  
ACCESSION AX082246  
VERSION AX082246.1 GI:13171007  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 740)  
AUTHORS Bertin,J.  
TITLE Novel molecules of the card-related protein family and uses thereof  
JOURNAL Patent: WO 0100826-A 48 04-JAN-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source  
1..740  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
54..641  
/note="unnamed protein product"

CDS  
/codon\_start=1  
/protein\_id="CAC33147.1"  
/db\_xref="GI:13171008"  
/translation="MGRARDIIDLAEENLTAELKFKLLSYPLRGYRIPRGAL  
LSMDALDTDLKLVSYLETYGAEITANVLRDMGLQEMAGQLQAAIHQSGAAPGIA  
PPOSAAKPGELHIDQRAALIARVNEMWLDALYGVLTDRQYAVABEPTNSKMR  
KLFSFTPANMWTCKDLLQALRESQSYLVEIDERS"

BASE COUNT 146 a 238 c 236 g 120 t

ORIGIN

Query Match 51.4%; Score 322; DB 6; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.6e-34;  
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 CGGCTGACAGCGGGGTGAGCGCGGCGCGGATCTGAGACCATGGGCGCGCC 169  
||| |  
Db 7 CGGCTGACAGCGGGGTGAGCGCGGCGCGGATCTGAGACCATGGGCGCGCC 66  
||| |  
QY 170 GCGAGCCATCTCTGATGCGCTGAGAGACCTGACCGCGGAGAGACTCAAGACTTCAGAC 229  
||| |



Db 67 GCAGCCATCCGTGATGCGCTGGAAGAACCTGACCGCGAGAGACTCAAAAGTTTCAAGC 126  
QY 230 TGAAGCTGCTGTCGATGCGCGCTGCGAGAGGCTACGGGGCGATCCCGGGGGCGCTGCG 289  
Db 127 TGAAGCTGCTGTCGATGCGCGCTGCGAGAGGCTACGGGGCGATCCCGGGGGCGCTGCG 186  
QY 290 TGTCCATGAGAGGCTTGGACCTGACGACAGCTGCTGACGCTTCTACCTGAGAGACTAGC 349  
Db 187 TGTCCATGAGAGGCTTGGACCTGACGACAGCTGCTGACGCTTCTACCTGAGAGACTAGC 246  
QY 350 GCGCCGAGCTCAACGCTTACGCTGCGCGAGACTGCGGCTTCAAGAGACTGCGGGCGAGC 409  
Db 247 GCGCCGAGCTCAACGCTTACGCTGCGCGAGACTGCGGCTTCAAGAGACTGCGGGCGAGC 306  
QY 410 TGCAGGCGGCGCAGCAGCAGG 431  
Db 307 TGCAGGCGGCGCAGCAGCAGG 328

RESULT 15  
AF384665 740 bp mRNA linear PRI 19-JUN-2001  
LOCUS  
DEFINITION Homo sapiens caspase recruitment domain protein 5 mRNA, complete cds.  
ACCESSION AF384665  
VERSION AF384665.1 GI:14488058  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Berlin, J.  
TITLE CARD5 protein is a CARD/PYRIN family member that is involved in apoptosis signal transduction  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 740)  
AUTHORS Berlin, J.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAY-2001) Neurobiology, Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA  
FEATURES  
source  
1..740  
location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
54..641  
/note="CARD5; CARD/PYRIN family member; bipartite protein comprised of an N-terminal PYRIN domain and a C-terminal CARD domain; involved in apoptosis signal transduction"  
/codon\_start=1  
/product="caspase recruitment domain protein 5"  
/protein\_id="AAK63850.1"  
/db\_xref="GI:14488058"  
/translation="MGARDAILDALLENLTAEIKKKFKLKLSPILREYGRIPRGAL  
LSMDALDTKIVSYLETYGAELTANVLRDMGLQEMAGQLAATHQSGAAPAGIA  
PPQSAKPGIHLFIIDHRAALILRVINVMILDALYGVILDEOYQAVRAEPTNPSKMR  
KLFSFTPAWMWTCCKDILLOALRESQSYLVEDLERS"

## CDS

Query Match 51.44; Score 322; DB 9; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.ee-34;  
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 110 CGGCTCAGCAGGGGTGAGCGGCGAGCGCGGCGGATCTGAGCCATGCGGCGCGCGC 169  
Db 7 CGGCTCAGCAGGGGTGAGCGGCGAGCGCGGCGGATCTGAGCCATGCGGCGCGCGCGC 66  
QY 170 GCGAGCGCATCCTGATGCTGAGAGAACTGACCGCGAGAGACTCAAGAAGTTCAAGC 229  
Db 67 GCGAGCGCATCCTGATGCTGAGAGAACTGACCGCGAGAGACTCAAGAAGTTCAAGC 126

QY 230 TGAAGCTGCTGTCGATGCGCGCTGCGAGAGGCTACGGGGCGATCCCGGGGGCGCTGCG 289  
Db 127 TGAAGCTGCTGTCGATGCGCGCTGCGAGAGGCTACGGGGCGATCCCGGGGGCGCTGCG 186  
QY 290 TGTCCATGAGAGGCTTGGACCTGACGACAGCTGCTGACGCTTCTACCTGAGAGACTAGC 349  
Db 187 TGTCCATGAGAGGCTTGGACCTGACGACAGCTGCTGACGCTTCTACCTGAGAGACTAGC 246  
QY 350 GCGCCGAGCTCAACGCTTACGCTGCGCGAGACTGCGGCTTCAAGAGACTGCGGGCGAGC 409  
Db 247 GCGCCGAGCTCAACGCTTACGCTGCGCGAGACTGCGGCTTCAAGAGACTGCGGGCGAGC 306  
QY 410 TGCAGGCGGCGCAGCAGCAGG 431  
Db 307 TGCAGGCGGCGCAGCAGCAGG 328

Search completed: December 22, 2002, 06:13:34  
Job time : 27:50 secs

...

\_\_\_\_\_

\_\_\_\_\_

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:21:01 ; Search time 257 Seconds  
(without alignments)  
5485.412 Million cell updates/sec

Title: US-09-691-763B-4

Perfect score: 626  
Sequence: 1 agcgcccttgcgtgagggc.....ggaagggaaggatcatt 626

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: N\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	626	100.0	22 AAD03891	CPG island of huma
2	626	100.0	22 AAD03889	Human target of me
3	389	62.1	22 AAH34052	Human colon cancer
4	389	62.1	21 AAC77884	Human cancer assoc
5	368	58.8	20 AA33631	Human breast tumou
6	361	57.7	24 AA447126	Pylin domain conta
7	342.6	54.8	22 AAD03806	Alternatively spli
8	342.6	54.7	22 AAD03907	Alternatively spli
9			22 AAD03890	Human target of me

10	340	54.3	340	22 AAD03892	Human target of me
11	322	51.4	740	22 AAF30007	Human CARD-5 cDNA
12	322	51.4	740	24 ABK87966	Human caspase recr
13	322	51.4	740	24 ABK87967	Human caspase recr
14	321	51.3	745	22 AAS59817	Human novel cytoK1
15	320	51.1	432	21 AAC98638	Human colon cancer
16	178.8	28.6	777	22 AAF30008	Mouse CARD-5 cDNA
17	178.8	28.6	777	24 ABK87964	Mouse caspase recr
18	178.8	28.6	777	24 ABK87965	Mouse caspase recr
19	178.6	28.5	803	22 AAD03904	Mouse target of me
20	177.6	28.4	551	24 AA447124	Pylin domain conta
21	174.4	27.9	807	22 AAC88099	Human FLEXHT-30 nu
22	143	22.8	639	22 AAH34053	Human colon cancer
23	117.8	18.8	605	22 AAD03905	Rat target of meth
24	98.6	15.8	180	22 AAS59860	Human novel cytoK1
25	86.2	13.8	16891	20 AAX37084	MEFV gene sequence
26	74.8	11.9	3500	24 AA447125	Pylin domain conta
27	74.8	11.9	3512	20 AAX37085	PMF associated pro
28	74	11.8	394	22 ABA50929	Human breast cell
29	74	11.8	394	22 ABA58906	Human foetal liver
30	74	11.8	394	22 ABA58959	Probe #14325 for g
31	74	11.8	394	22 AAK17239	Human brain expies
32	74	11.8	394	22 AAK43026	Human bone marrow
33	74	11.8	394	22 AA123791	Probe #13724 for g
34	74	11.8	394	22 AA149102	Probe #17788 used t
35	74	11.8	394	22 AA109399	Probe #9390 used t
36	74	11.8	394	22 ABA37092	Human genome-deriv
37	68.4	10.9	365	22 AAD08836	Human G-protein co
38	68.4	10.9	3536	24 AA447132	Pylin domain conta
39	68.4	10.9	6461	22 AAS00033	Human ATRAS-2-enco
40	62.8	10.0	368	22 ABA45810	Human breast cell
41	62.8	10.0	368	22 ABA56323	Human foetal liver
42	62.8	10.0	368	22 ABA25964	Probe #4430 for ge
43	62.8	10.0	368	22 AAK04503	Human brain expies
44	62.8	10.0	368	22 AAK30003	Human bone marrow
45	62.8	10.0	368	22 AA114595	Probe #4528 for ge

#### ALIGNMENTS

RESULT 1	
AAD03891	
ID	AAD03891 standard; DNA; 626 BP.
XX	AC
XX	AAD03891;
DT	02-JUL-2001 (first entry)
XX	
DE	CPG island of human TMS1 gene.
XX	
KW	Human: target of methylation-induced silencing-1; TMS1; cytostatic;
KW	antiproliferative; apoptosis inducer; gene therapy; Cpg island;
KW	caspase-recruiting domain; CARD; cancer; breast; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200129235-A2.
XX	
PD	26-APR-2001.
XX	
PF	18-OCT-2000; 2000MO-US28747.
XX	
PR	18-OCT-1999; 99US-0159975.
XX	
PA	(UYEM-) UNIV EMORY.
XX	
PI	Verlino PW;
XX	
DR	WPI; 2001-290922/30.
XX	
PT	Novel gene TMS1, transcriptionally silenced due to increased
PT	methylation useful for identifying subject at risk of developing tumor





RESULT 4  
AAC77884  
ID AAC77884 standard; cDNA; 811 BP.  
XX  
AC AAC77884;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human cancer associated gene sequence SEQ ID NO:278.  
XX  
KW Human; cancer associated gene; cancer antigen; detection; cancer;  
diagnosis; cytostatic; proliferative; vulnery; immunomodulator;  
antidiabetic; antiaesthetic; antirheumatic; antibacterial; antiviral;  
antiinflammatory; antihypoid; antiallergic; antibacterial; cardiant;  
dermatological; neuroprotective; thrombolytic; coagulant; neotropic;  
vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;  
immune disorder; haematopoietic cell disorder; autoimmune disorder;  
allergic reaction; graft versus host disease; organ rejection;  
haemostatic; thrombolytic; cardiovascular disorder; infection;  
neurological disease; drug screening; ss.  
XX  
OS Homo sapiens.  
XX  
PN M0200055350-A1.  
XX  
PD 21-SEP-2000.  
XX  
PE 08-MAR-2000; 2000M0-0S05882.  
XX  
PR 12-MAR-1999; 990S-0124270.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI: 2000-587533/55.  
XX  
DR P-PSDB; AAB43675.  
XX  
PT Novel isolated nucleic acids comprising sequences encoding peptides  
PT useful for treating or diagnosing e.g. cancer -  
XX  
XX  
PS Claim 1; Page 841; 2352bp; English.  
XX  
CC AAC77607 to AAC78448 encode the human cancer associated proteins given  
CC in AAB43398 to AAB44239. The proteins can have activities based on the  
CC tissues and cells the genes are expressed in. Example of activities  
CC include: cytostatic; proliferative; vulnery; immunomodulator;  
CC antidiabetic; antiaesthetic; antirheumatic; antibacterial;  
CC antiinflammatory; antihypoid; antiallergic; antibacterial;  
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
CC neotropic; vasotropic; antiporiatic and antiangiogenic. The  
CC polynucleotides and polypeptides can be used for preventing, treating or  
CC ameliorating medical conditions and diagnosing pathological conditions.  
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
CC the present invention may be used to treat immune disorders by activating  
CC or inhibiting the proliferation, differentiation or mobilisation of  
CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
CC disorders, allergic reactions, graft versus host disease and organ  
CC rejection, modulate haemostatic or thrombolytic activity, modulate  
CC inflammation, cancers, cardiovascular disorders, neurological disease and  
CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
CC the present invention.  
XX  
SO Sequence 811 BP; 170 A; 253 C; 259 G; 128 T; 1 other;

Query Match 62.1%; Score 389; DB 21; Length 811;  
Best Local Similarity 100.0%; Pred. No. 3.2e-61;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

43 CGAGACCAAGTGGAGAGAGCGGAGTCACAGTTCGCCGCCGAGCCAGCCACTTCTCC 102

DB 1 GAGAGACAGAGTGGAGAGAGAGCGGAGAGTCCAGGTTCCGCCCGAGCCGACTTCTCC 60  
QY 103 TGGTCGGCGCTGAGAGCGGGGTGAGCGGCGAGCGGCCGGGATCTGGAGGCATGGG 162  
DB 61 TGGTCGGCGCTGAGAGCGGGGTGAGCGGCGAGCGGCCGGGATCTGGAGGCATGGG 120  
QY 163 CGGCGCGCGGAGCGCATCTGAGATGGGCTGAGAACTGACCGCGGAGAGCTCAAGAG 222  
DB 121 CGGCGCGCGGAGCGCATCTGAGATGGGCTGAGAACTGACCGCGGAGAGCTCAAGAG 180  
QY 223 TTCAAGCTGAAGCTGCTGTGCGTCCGCTGCGCGAGGCTCAAGGCGCATCCCGCGGC 282  
DB 181 TTCAAGCTGAAGCTGCTGTGCGTCCGCTGCGCGAGGCTCAAGGCGCATCCCGCGGC 240  
QY 283 GCGCTGCTGTGCTGAGAGCGGCTTGGAGCTACCGAGAGCTGTACCTTCTACCTGGAG 342  
DB 241 GCGCTGCTGTGCTGAGAGCGGCTTGGAGCTACCGAGAGCTGTACCTTCTACCTGGAG 300  
QY 343 ACCTACGGCGCGGAGCTACCGCTTAACTGTCTGCGCGAGCATGGGCTGCGAGATGGCC 402  
DB 301 ACCTACGGCGCGGAGCTACCGCTTAACTGTCTGCGCGAGCATGGGCTGCGAGATGGCC 360  
QY 403 GCGCAGCTGAGCGCGCGCGCGAGCGAGG 431  
DB 361 GCGCAGCTGAGCGCGCGCGCGAGCGAGG 389

RESULT 5  
AAZ3631  
ID AAZ3631 standard; cDNA; 779 BP.  
XX  
AC AAZ3631;  
XX  
DT 08-DEC-1999 (first entry)  
XX  
DE Human breast tumour-associated EST 21.  
XX  
KW Expressed sequence tag; EST; human; breast; cancer; gene therapy;  
KW treatment; tumour; cytostatic; medicament; ss.  
XX  
OS Homo sapiens.  
XX  
PN DE19813839-A1.  
XX  
PD 23-SEP-1999.  
XX  
PE 20-MAR-1998; 98DE-1013839.  
XX  
PR 20-MAR-1998; 98DE-1013839.  
XX  
PA (META-) METACEN GES GENOMFORSCHUNG MBH.  
XX  
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;  
XX WPI: 1999-528981/45.  
XX  
PT Human nucleic acid sequences and protein products from tumor breast  
PT tissue, useful for breast cancer therapy -  
XX  
XX  
PS Claim 1a; 101; 188bp; German.  
XX  
CC This invention describes novel human nucleic acid sequences from tumor  
CC breast tissue which have cytostatic activity. The nucleic acid sequences  
CC can be used to produce and isolate full-length gene sequences. They can  
CC be used to express proteins, which can be used as tools to find an  
CC activity against breast cancer. The sequences can be used in sense or  
CC antisense form. They are especially useful for medicaments for gene  
CC therapy to treat breast cancer. AAZ3611-748617 represents expressed  
CC sequence tags described in the method of the invention.  
XX  
SO Sequence 779 BP; 149 A; 252 C; 251 G; 127 T; 0 other;

Query Match 58.8%; Score 368; DB 20; Length 779;  
Best Local Similarity 100.0%; Pred. No. 1.8e-57;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 64 GCGGGAGTCCAGGTTCCGCCCCGAGCCGACTTCTCTGTCGGCGGCTGCACGGGG 123
    |||||||
DB 1 GCGGGAGTCCAGGTTCCGCCCCGAGCCGACTTCTCTGTCGGCGGCTGCACGGGG 60
OY 124 TGAGCGCGCGAGCGCGCGGATCCTGAGCCCAATGGGGCGCGCGGAGCCATCCTG 183
    |||||||
DB 61 TGAGCGCGCGAGCGCGCGGATCCTGAGCCCAATGGGGCGCGCGGAGCCATCCTG 120
OY 184 GATGCGCTGGAGAACTGACCGCCGAGAGCTCAAGAATTCAAGTGAAGTCTGTCG 243
    |||||||
DB 121 GATGCGCTGGAGAACTGACCGCCGAGAGCTCAAGAATTCAAGTGAAGTCTGTCG 180
OY 244 GTGCGCGTGCAGAGGCTGACGGGCGCAATCCGCGGGCGCGCTGTCGTCATGAGACCC 303
    |||||||
DB 181 GTGCGCGTGCAGAGGCTGACGGGCGCAATCCGCGGGCGCGCTGTCGTCATGAGACCC 240
OY 304 TTGAGACCTACCCGACAAAGCTGTCAGCTTCTACTGGAGACTACGGCGCCGACTCACC 363
    |||||||
DB 241 TTGAGACCTACCCGACAAAGCTGTCAGCTTCTACTGGAGACTACGGCGCCGACTCACC 300
OY 364 GCTAACGTGCTGCAGCATGGGCTGCGAGAGATGGCGGGGAGGCTGAGGCGGCGCAGG 423
    |||||||
DB 301 GCTAACGTGCTGCAGCATGGGCTGCGAGAGATGGCGGGGAGGCTGAGGCGGCGCAGG 360
OY 424 CACCGAGG 431
    |||||||
DB 361 CACCGAGG 368
```

RESULT 6  
ID AAL47126  
XX AAL47126 standard; DNA; 782 BP.

AC AAL47126;

DT 20-AUG-2002 (first entry)

XX Pycin domain containing protein Pycard coding sequence.

DE Pycin domain: PYD domain; antiinflammatory; antiparkinsonian;  
XX antiarteriosclerotic; antispasmodic; antibacterial; vinuicide;  
KW neuroprotective; antiarthritic; antirheumatic; antileptostomat;  
KW nephrotoxic; osteopathic; noctropic; intracellular signal transduction;  
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;  
KW osteoarthritis; glomerulonephritis; gene; ds.

XX Unidentified.

OS WO200240668-A2.

PN 23-MAY-2002.

PD 30-OCT-2001; 2001WO-EP12545.

PF 15-NOV-2000; 2000DE-1056687.

PR 30-NOV-2000; 2000DE-1059595.

XX (APOT-) APOTEC RES & DEV LTD.

XX Tschoopp J, Martignon F;

XX WPI: 2002-427093/45.

DR P-PSDB: AAO17854.

XX New DNA encoding protein with pyrin domain, useful for treating  
PT diseases involving impaired signal transduction, particularly  
PT inflammation, also proteins and antibodies -  
XX

PS Claim 5; Fig 1; 11pp; German.

XX The present invention relates the DNA and their encoded proteins, where  
CC the proteins constrain at least one PYD (pyrin) domain. These can be used  
CC to treat diseases associated with impaired intracellular signal  
CC transduction, particularly inflammation such as psoriasis,  
CC arteriosclerosis, bacterial or viral infections (particularly meningitis  
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,  
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's  
CC and Parkinson's diseases. The present sequence is a coding sequence of  
CC the invention.

XX Sequence 782 BP; 158 A; 251 C; 246 G; 127 T; 0 other.

Query Match 57.7%; Score 361; DB 24; Length 782;  
Best Local Similarity 100.0%; Pred. No. 3.3e-56;  
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 71 GTCCAGGTTCCGCCCCGAGCGGAGCTTCTCTGTCGGCGGCTGCACGGGGTGAAGCG 130
    |||||||
DB 1 GTCCAGGTTCCGCCCCGAGCGGAGCTTCTCTGTCGGCGGCTGCACGGGGTGAAGCG 60
OY 131 CCGCAGCGCGCGGGGATCCCTGAGACCAATGGGCGCGCGGAGCCATCCTGATGCGC 190
    |||||||
DB 61 CCGCAGCGCGCGGGGATCCCTGAGACCAATGGGCGCGCGGAGCCATCCTGATGCGC 120
OY 191 TGAGAACTGACCGCGCGAGGCTCAAGAAGTTCAAGCTGAAGCTGCTGCGATGCGCG 250
    |||||||
DB 121 TGAGAACTGACCGCGCGAGGCTCAAGAAGTTCAAGCTGAAGCTGCTGCGATGCGCG 180
OY 251 TGGCGAGGCTTACGGGCGCATCCCGGGGGCGCGCTGCTGTCATGAGCCCTTGGAGC 310
    |||||||
DB 181 TGGCGAGGCTTACGGGCGCATCCCGGGGGCGCGCTGCTGTCATGAGCCCTTGGAGC 240
OY 311 TCACCGAAGCTGTCAGCTTCTACTGAGACTACGGCGCGAGCTCACCGCTTACG 370
    |||||||
DB 241 TCACCGAAGCTGTCAGCTTCTACTGAGACTACGGCGCGAGCTCACCGCTTACG 300
OY 371 TGCTGCGCGACATGGGCTGCGAGAGATGGCGGCGAGCTGCGAGCGGCGCAGCAGG 430
    |||||||
DB 301 TGCTGCGCGACATGGGCTGCGAGAGATGGCGGCGAGCTGCGAGCGGCGCAGCAGG 360
OY 431 G 431
    |
DB 361 G 361
```

RESULT 7  
ID AAD03906  
XX AAD03906 standard; CDNA; 713 BP.

AC AAD03906;

DT 02-JUL-2001 (first entry)

XX Alternatively spliced form of human TMS1 cDNA (lacking exon2).

DE Human; target of methylation-induced silencing-1; TMS1; cytoskeletal;  
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;  
KW caspase-recruiting domain; CARD; cancer; breast; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 75..605

FT /tag= a /product= "Alternatively spliced form of human TMS1  
FT protein lacking exon2"

PN WO200129235-A2.

XX 26-APR-2001.

[illegible]

DE	Alternatively spliced form of human TMS1 cDNA (lacking exon3).	
XX		
KM	Human: target of methylation-induced silencing-1; TMS1; cytostatic;	
KM	antiapoptotic; apoptosis inducer; gene therapy; CpG island;	
KM	caspase-recruiting domain; CARO; cancer; breast; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	75..404
FT		/*tag= a
FT		/product= "Alternatively spliced form of human TMS1
FT		protein lacking exon3"
FT		/note= "CDS does not include stop codon"
XX		/partial
XX		
PN	W0200129235-A2.	
XX		
PD	26-APR-2001.	
XX		
XX	18-OCT-2000; 2000MO-US28747.	
PF		
XX		
PR	18-OCT-1999; 99US-0159975.	
XX		
PA	(UTEM-) UNIV EMORY.	
PI	Vertino PM;	
XX		
DR	WPI; 2001-290922/30.	
XX		
DR	P-PSDB; AAE0595.	
XX		
PT	Novel gene TMS1, transcriptionally silenced due to increased	
PT	methylation useful for identifying subject at risk of developing tumor	
PT	characterized by abnormal methylation, for treating cancer by inducing	
PT	apoptosis	
XX		
XX		
XX	Claim 106; Page 123-124; 124pp: English.	
XX		
CC	The invention relates to identification of target of methylation-induced	
CC	silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to	
CC	abnormal methylation of a CpG island in its 5' regulatory region. TMS1	
CC	consists of a carboxy terminal caspase-recruiting domain (CARO) and	
CC	plays a role in induction of apoptosis. TMS1 gene and protein are useful	
CC	as tools for diagnosing and treating a subject at risk of developing	
CC	cancer (e.g. breast cancer) characterized by abnormal CpG methylation or	
CC	abnormally low levels of TMS1 expression products. Unique fragments of	
CC	TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.	
CC	Increasing TMS1 polypeptide level to an above normal level. The CpG	
CC	island region of TMS1 or its fragments are used to study the methylation	
CC	patterns apart from any coding region contained in it.	
CC	The present sequence is alternatively spliced form of human target of	
CC	methylation-induced silencing-1 (TMS1) cDNA lacking exon3.	
CC	Note: The present sequence designated as SRO ID NO:26, is referred	
CC	as DNA throughout the specification, however this sequence has been	
CC	referred as protein in claim 106.	
XX		
SO	Sequence 405 BP; 63 A; 139 C; 147 G; 56 T; 0 other:	
XX		
Query Match	54.7%; Score 342.6; DB 22; Length 405;	
Best Local Similarity	98.9%; Pred. No. 6.ee-53;	
Matches 345; Conservative	0; Mismatches 4; Indels 0; Gaps	0;
OY	83 CCCCAGGAGCGACGCTTCCTCCGCTGGTGGGGGCGTGCAGCGGGGTAGGGGGGCGAGCGGCGG	142
DB	1 CACAGCGTCCGACCTTCTCCTGCTGGCGGCGTGCAGCGGGGTAGCGCGGCGAGCGGCGG	60
OY	143 GGGATCTTGGAGCCATGCGGGCGCGCGCGAGCGCCATCTTGGATGCGCTGGAGAACTGGA	202
DB	61 GGGATCTTGGAGCCATGCGGGCGCGCGCGAGCGCCATCTTGGATGCGCTGGAGAACTGGA	120
OY	203 CCCCAGGAGCGACGCTTCAAGATTCAACCTTCAACCTCTGCTGGTGGCGCGCGAGCGGACT	262



Db 121 CCGCGGAGAGCTCAAGATTCAAGCTGAAGCTGCTGTGCGGTGCGCGGAGGGCT 180  
Qy 263 ACGGGCGCATCCCGCGGGGCGCGTGTCTGTCTCATGAGACCTTGGACTCAGCACAAGC 322  
Db 181 ACGGGGCGCATCCCGGGGCGCGTGTCTGTCTCATGAGACCTTGGACTCAGCACAAGC 240  
Qy 323 TGGTCAGCTTCTACCTGAGAGACTACGGCGCGAGCTCAACGCTTAACGTGTGCGGACCA 382  
Db 241 TGGTCAGCTTCTACCTGAGAGACTACGGCGCGAGCTCAACGCTTAACGTGTGCGGACCA 300  
Qy 383 TGGGCGCTGAGAGAGATGGCGCGGCGAGCTGCAAGCGCGCCACGACCAAGG 431  
Db 301 TGGGCGCTGAGAGAGATGGCGCGGCGAGCTGCAAGCGCGCCACGACCAAGG 349

RESULT 9  
AAD03890 standard; cDNA: 770 BP.

XX AC AAD03890;  
XX DT 02-JUL-2001 (first entry)

XX DE Human target of methylation-induced silencing-1 (TMS1) cDNA.

XX KW Human: target of methylation-induced silencing-1; TMS1: cytosstatic;  
XX KM caspase-recruiting domain; CARD; cancer; breast; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT CDS 75..662  
XX FT /\*tag= a  
XX FT /product= "Human TMS1 protein"

XX PN WO200129235-A2.

XX PD 26-APR-2001.

XX PF 18-OCT-2000; 2000WO-US28747.

XX PR 18-OCT-1999; 99US-0159975.

XX PA (UYEM-) UNIV EMORY.

XX PI Vertino PM;

XX DR WPI: 2001-290922/30.

XX DR P-PSDB; AAE00588.

XX PT Novel gene TMS1, transcriptionally silenced due to increased  
XX PT methylation useful for identifying subject at risk of developing tumor  
XX PT characterized by abnormal methylation, for treating cancer by inducing  
XX PT apoptosis

XX PS Claim 68; Page 113-114; 124pp; English.

XX The invention relates to identification of target of methylation-induced  
XX silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to  
XX abnormal methylation of a CpG island in its 5' regulatory region. TMS1  
XX consists of a carboxy terminal caspase-recruiting domain (CARD) and  
XX plays a role in induction of apoptosis. TMS1 gene and protein are useful  
XX as tools for diagnosing and treating a subject at risk of developing  
XX cancer (e.g. breast cancer) characterized by abnormal CpG methylation or  
XX abnormally low levels of TMS1 expression products. Unique fragments of  
XX TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.  
XX TMS1 molecule is also useful for treating abnormal cell proliferation by  
XX increasing TMS1 polypeptide level to an above normal level. The CpG  
XX island region of TMS1 or its fragments are used to study the methylation  
XX patterns apart from any coding region contained in it.  
XX The present sequence is human target of methylation-induced silencing-1  
XX (TMS1) cDNA.

XX SQ Sequence 770 BP; 157 A; 247 C; 241 G; 125 T; 0 other;

XX Query Match 54.7%; Score 342.6; DB 22; Length 770;

XX Best Local Similarity 98.9%; Pred. No. 6.5e-53;

XX Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 83 CCCGAGACGCACTTCTCTCTGTCGGCGGCTGACAGCGGGGTAGAGCGCGAGCGCG 142

Db 1 CCACGCGTCCACTTCTCTCTCTGTCGGCGGCTGACAGCGGGGTAGAGCGCGAGCGCG 60

Qy 143 GGGATCTTGAAGCCATGAGGCGCGCGCGGACCCATCTTGATGCGCTGAGAACTGA 202

Db 61 GGGATCTTGAAGCCATGAGGCGCGCGCGGACCCATCTTGATGCGCTGAGAACTGA 120

Qy 203 CCCGAGAGAGCTCAAGATTCAAGCTGAAGCTGCTGAGTCCGCTGCGGAGGGCT 262

Db 121 CCGCGGAGAGACTCAAGATTCAAGCTGAAGCTGCTGAGTCCGCTGCGGAGGGCT 180

Qy 263 ACGGGCGCATCCCGGGGCGCGTGTCTGTCTCATGAGACCTTGGACTCAGCACAAGC 322

Db 181 ACGGGCGCATCCCGGGGCGCGTGTCTGTCTCATGAGACCTTGGACTCAGCACAAGC 240

Qy 323 TGGTCAGCTTCTACCTGAGAGACTACGGCGCGAGCTCAACGCTTAACGTGTGCGGACA 382

Db 241 TGGTCAGCTTCTACCTGAGAGACTACGGCGCGAGCTCAACGCTTAACGTGTGCGGACA 300

Qy 383 TGGGCGCTGAGAGAGATGGCGCGGCGAGCTGCAAGCGCGCCACGACCAAGG 431

Db 301 TGGGCGCTGAGAGAGATGGCGCGGCGAGCTGCAAGCGCGCCACGACCAAGG 349

RESULT 10  
AAD03892 standard; DNA: 340 BP.

XX AC AAD03892;  
XX DT 02-JUL-2001 (first entry)

XX DE Human target of methylation-induced silencing-1 (TMS1) exon1 DNA.

XX KW Human: target of methylation-induced silencing-1; TMS1: cytosstatic;  
XX KM antiproliferative; apoptosis inducer; gene therapy; CpG island;  
XX KM caspase-recruiting domain; CARD; cancer; breast; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT CDS 67..339  
XX FT /\*tag= a  
XX FT /product= "Human TMS1 exon1 protein"  
XX FT /note= "CDS does not include stop codon"  
XX FT /partial

XX PN WO200129235-A2.

XX PD 26-APR-2001.

XX PF 18-OCT-2000; 2000WO-US28747.

XX PR 18-OCT-1999; 99US-0159975.

XX PA (UYEM-) UNIV EMORY.

XX PI Vertino PM;

XX DR WPI: 2001-290922/30.

XX DR P-PSDB; AAE00589.

XX PT Novel gene TMS1, transcriptionally silenced due to increased  
XX PT methylation useful for identifying subject at risk of developing tumor  
XX PT characterized by abnormal methylation, for treating cancer by inducing

PT	apoptosis -
XX	
PS	Claim 68; Page 115; 124pp; English
US	

The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression products. Unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy. TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it. The present sequence is human target of methylation-induced silencing-1 (TMS1) exon1 DNA.

Sequence 340 BP; 53 A; 112 C; 127 G; 48 T; 0 other;

Query Match 54.3%; Score 340; DB 22; Length 340;  
Best Local Similarity 100.0%; Pred. No. 1.9e-52;  
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0

[illegible]

RESULT 11  
AAAF30007  
ID    AAAF30007    standard; CDNA; 740 BP.

AC      AAF30007;  
XX  
DT      23-APR-2001 (first entry)

Human CARD-5 cDNA

KW CARP-5, caspase recruitment domain; human; cancer; infection;  
 KW autoimmune disease; neurological disease; haematological disease  
 KW immune disease; inflammation; antitumour; antiseptic;  
 KW immunomodulator; antinflammatory; apoptosis; diagnosis;  
 KW gene therapy; ss.

OS Homo sapiens.

Key	Location/Qualifiers
FH	54..641
FT	/lag- a
FT	/note- "the open reading frame is also specifically
FT	

Db	307	TGCAAGCGCCACGACGACGAGG	328
RESULT 12			
ID	ABK87966		
XX	ABK87966	standard; cDNA; 740 BP.	
AC	ABK87966;		
DT	07-OCT-2002	(first entry)	
XX			
DE	Human caspase recruitment domain-5 (CARD-5) cDNA.		
KW	Human; gene; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;		
KW	immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;		
KW	apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;		
KW	cell proliferation; gene therapy; immune disorder;		
KW	chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;		
KW	sarcoidosis; atopy; asthma; allergy; glomerular nephritis;		
KW	human immunodeficiency virus; HIV; bacterial infection; tuberculosis;		
KW	lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;		
KW	arthritis; cell depletion; neurological disorder; Alzheimer's disease;		
KW	Parkinson's disease; spinal muscular atrophy; haematologic disease;		
KW	myelodysplastic syndrome; aplastic anaemia; myocardial infarction;		
KW	stroke.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	54..641	
FT		/*tag" a	
FT		/product= "CARD-5"	
XX			
PN	W0200244354-A2.		
XX			
PD	06-JUN-2002.		
XX			
PF	29-NOV-2001; 2001WO-US44894.		
XX			
PR	01-DEC-2000; 2000US-0728721.		
PR	24-APR-2001; 2001US-0841879.		
XX			
PA	(MILL-) MILLENIUM PHARM INC.		
XX			
PI	Berlin J;		
XX			
DR	WPI; 2002-557538/59.		
DR	P-PSDB; AA099353.		
PT			
PT	Novel isolated murine or human caspase recruitment domain (CARD)-5		
PT	polypeptide, useful for treating immune disorders such as Hashimoto's		
PT	thyroiditis, graft rejection, allergy, glomerular nephritis,		
PT	tuberculosis		
XX			
PS	Claim 7; Fig 3; 100pp; English.		
XX			
CC	The invention discloses the isolated polypeptides, and encoding nucleic		
CC	acids, of murine and human caspase recruitment domain (CARD)-5. Caspases		
CC	(cysteineyl aspartate-specific proteinases) are central to the apoptotic		
CC	program and responsible for the degradation of cellular proteins that		
CC	lead to the morphological changes seen in cells undergoing apoptosis.		
CC	Caspases interact with other caspases via their CARDs and different		
CC	subtypes of CARDs may confer binding specificity. CARD-5 is an		
CC	intracellular protein that is predicted to be involved in regulating		
CC	caspase activation. CARD-5 activates the nuclear factor-kappa B		
CC	(NF-kappaB) transcription factor pathway and binds the CARDs of		
CC	caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5		
CC	activity and NF-kappaB activation, regulate cell growth and cell death		
CC	and be used in gene therapy. The CARD-5 polypeptides are useful for		
CC	identifying compounds which bind to them and modulate their activity and		
CC	for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,		
CC	nucleic acids, antibodies and modulators of CARD-5 expression or activity		
CC	can be used to treat immune disorders such as chronic inflammatory		

	CC	diseases and disorders, Hashimoto's thyroiditis, graft rejection,
	CC	sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
	CC	nephritis, human immunodeficiency virus (HIV) and bacterial infections
	CC	(including tuberculosis and lepromatous leprosy) and in screening and
	CC	detection assays. Modulators of CARD-5 activity or expression are also
	CC	useful for treating autoimmune disorders, such as systemic lupus
	CC	erythematosus and arthritis, cell depletion, neurological disorders,
	CC	such as Alzheimer's disease, Parkinson's disease and spinal muscular
	CC	atrophy, haematologic diseases, such as myelodysplastic syndrome and
	CC	aplastic anaemia, myocardial infarction and stroke. The sequence
	CC	presented is the human caspase recruitment domain-5 (CARD-5) cDNA.
XX	SQ	Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;
XX		
XX	Query Match	51.4%; Score 322; DB 24; Length 740;
XX	Best Local Similarity	100.0%; Pred. No. 3.2e-49;
XX	Matches 322;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	110	CGGCTGCAAGCGGGGTGAAGCGGCAGACGGCGCGGGATCCTGGAGCATGATGGCGCGCCG 169
DB	7	CGGTCGCAAGCGGGGTGAAGCGGCAGACGGCGCGGGATCCTGGAGCATGATGGCGCGCGC 66
QY	170	GCGAGCGCCATCTTGATGATGCGCTGGAGAACCTGACGCCCGAGAGAGCTCAAGAATTTCAMG 229
DB	67	GCGAGCGCCATCTTGATGATGCGCTGGAGAACCTGACGCCCGAGAGAGCTCAAGAATTTCAMG 126
QY	230	TGAAACTGCTGTCCGTGCGCGTCCGAGAGGGCTACAGGGCGCATCCCGCGGGCGCGTGC 289
DB	127	TGAAGCTGCTGTCCGTGCGCGTCCGAGAGGGCTACAGGGCGCATCCCGCGGGCGCGTGC 186
QY	290	TGTCGATGAGAGCGCTTGGACTCACGACGAAGCTGTCACGCTTTACTTACCTGGAGACTTACG 349
DB	187	TGTCGATGAGAGCGCTTGGACTCACGACGAAGCTGTCACGCTTTACTTACCTGGAGACTTACG 246
QY	350	GCGCGGACCTACCGCTTAACGTGTCGCGCAGACATGAGGGCTCGCAGAGATGAGCCCGGAGAC 409
DB	247	GCGCGGACCTACCGCTTAACGTGTCGCGCAGACATGAGGGCTCGCAGAGATGAGCCCGGAGAC 306
QY	410	TGCAGGCGGCCACGACACCAAGGG 431
DB	307	TGCAGGCGGCCACGACACCAAGGG 328
XX		
XX	RESULT 13	
XX	ABK87967/c	
ID	ABK87967	standard; DNA: 740 BP.
XX		
XX	ABK87967;	
DT	07-OCT-2002	(first entry)
DE		
KW	Human	caspase recruitment domain-5 (CARD-5) cDNA complementary strand.
KM	Human; ss;	caspase recruitment domain-5; CARD-5; antiinflammatory;
KM	Immunosuppressive;	caspase; cysteinyl aspartate-specific proteinase;
KM	apoptosis;	nuclear factor-kappa B; NF-kappaB; transcription factor;
KM	cell proliferation;	gene therapy; immune disorder;
KM	chronic inflammation;	disease; Hashimoto's thyroiditis; graft rejection;
KM	sarcoidosis;	atrophy; asthma; allergy; glomerular nephritis;
KM	human immunodeficiency virus;	HIV; bacterial infection; tuberculosis;
KM	lepromatous leprosy;	autoimmune disorder; systemic lupus erythematosus;
KM	arthritis;	cell depletion; neurological disorder; Alzheimer's disease;
KM	Parkinson's disease;	spinal muscular atrophy; haematologic disease;
KM	myelodysplastic syndrome;	aplastic anaemia; myocardial infarction;
KM	stroke.	
XX		
XX	Homo sapiens.	
OS		
XX	WO200244354-a2.	
XX		
XX	06-JUN-2002.	
XX		
XX	29-NOV-2001;	2001MO-US44894.
XX		

XX 01-DEC-2000; 2000US-0728721.  
PR 24-APR-2001; 2001US-0841879.  
XX (MILL-) MILLENIUM PHARM INC.  
XX  
PI Bertin J;  
XX WPI: 2002-557538/59.  
XX  
XX Novel isolated murine or human caspase recruitment domain (CARD)-5  
PT polypeptide, useful for treating immune disorders such as Hashimoto's  
PT thyroiditis, graft rejection, allergy, glomerular nephritis,  
PT tuberculosis  
XX  
XX Claim 8; Fig 3; 100pp; English.  
XX  
XX The invention discloses the isolated polypeptides, and encoding nucleic  
CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases  
CC (cysteinyl aspartate-specific proteases) are central to the apoptotic  
CC program and responsible for the degradation of cellular proteins that  
CC lead to the morphological changes seen in cells undergoing apoptosis.  
CC Caspases interact with other caspases via their CARDS and different  
CC subtypes of CARDS may confer binding specificity. CARD-5 is an  
CC intracellular protein that is predicted to be involved in regulating  
CC caspase activation. CARD-5 activates the nuclear factor-kappa B  
CC (NF-kappaB) transcription factor pathway and binds the CARDS of  
CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5  
CC activity and NF-kappaB activation, regulate cell growth and cell death  
CC and be used in gene therapy. The CARD-5 polypeptides are useful for  
CC identifying compounds which bind to them and modulate their activity and  
CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,  
CC nucleic acids, antibodies and modulators of CARD-5 expression or activity  
CC can be used to treat immune disorders such as chronic inflammatory  
CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,  
CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular  
CC nephritis, human immunodeficiency virus (HIV) and bacterial infections  
CC (including tuberculosis and lepromatous leprosy) and in screening and  
CC detection assays. Modulators of CARD-5 activity or expression are also  
CC useful for treating autoimmune disorders, such as systemic lupus  
CC erythematosus and arthritis, cell depletion, neurological disorders,  
CC such as Alzheimer's disease, Parkinson's disease and spinal muscular  
CC atrophy, haematologic diseases, such as myelodysplastic syndrome and  
CC aplastic anaemia, myocardial infarction and stroke. The sequence  
CC presented is the complementary strand to the human caspase recruitment  
CC domain-5 (CARD-5) cDNA.  
XX  
XX Sequence 740 BP; 120 A; 236 C; 238 G; 146 T; 0 other;

Query Match 51.4%; Score 322; DB 24; Length 740;  
Best Local Similarly 100.0%; Pred. No. 3.2e-49;  
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 CGCGTCAGCGGGGTGAGGCGGCGGCGGAGTCTGTGAGCCATGAGGCGCGGCG 169  
DB 734 CGGCTGCAGCGGGGTGAGGCGGCGGCGGCGGAGTCTGTGAGCCATGAGGCGCGGCG 675  
QY 170 GCGAGCCATCTTGATGCGGTGAGAACTGACCGCGGAGGCTCAAGATTCAAGC 229  
DB 674 GCGAGCCATCTTGATGCGGTGAGAACTGACCGCGGAGGCTCAAGATTCAAGC 615  
QY 220 TGAAGCTGCTGTGAGTCCGTCGCGGAGGCTACGCGGCGGCGGCGGCGGCTGC 289  
DB 614 TGAAGCTGCTGTGAGTCCGTCGCGGAGGCTACGCGGCGGCGGCGGCGGCTGC 555  
QY 290 TGTCCATGAGAGCGCTTGAGACTTACCGCAAGCTGCTGACCTTCTACCTGAGAGCTTACG 349  
DB 554 TGTCCATGAGAGCGCTTGAGACTTACCGCAAGCTGCTGACCTTCTACCTGAGAGCTTACG 495  
QY 350 GCGCGAGGCTACCGCTAAGCTGTCGCGGAGCATGAGGCTGCGAGGAGATGAGCGCGGCGAGC 409  
DB 494 GCGCGAGGCTACCGCTAAGCTGTCGCGGAGCATGAGGCTGCGAGGAGATGAGCGCGGCGAGC 435

OY 410 TGCAGCGGCGCACGACCAAGG 431  
DB 434 TGCAGCGGCGCACGACCAAGG 413

RESULT 14  
AAS59817  
ID AAS59817 standard; cDNA; 745 BP.  
XX  
XX AAS59817;  
XX  
XX 16-JAN-2002 (first entry)  
XX  
XX Human novel cytokine encoding cDNA 790CIP2B\_1 #1.  
XX  
XX Human; ss; cytokine; cell proliferation; cell differentiation;  
KW antiinflammatory; stem cell growth factor; activin; inhibin; cancer;  
KW nervous system disease; neuropathy; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; spinal cord disorder;  
KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;  
KW platelet disorder; thrombocytopenia; stem cell disorder;  
KW aplastic anaemia; tissue regeneration; wound healing; ulcer;  
KW osteoporosis; osteoarthritis; bone degenerative disorder;  
KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;  
KW severe combined immunodeficiency; infection; autoimmune disorder;  
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;  
KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;  
KW inflammatory bowel disease; food supplement.

Homo sapiens.

W0200175093-AL.

11-OCT-2001.

30-MAR-2001; 2001WO-US10484.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

22-SEP-2000; 2000US-0668680.

23-OCT-2000; 2000US-0695618.

30-NOV-2000; 2000US-0728711.

14-MAR-2001; 2000US-0728711.

(HYSEQ-) HYSEQ INC.

Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;  
PI Yang Y, Zabo QA, Chen R, Wang D, Goodrich RW, Liu C, Dmanac RT;  
XX  
XX WPI: 2001-626432/72.  
XX  
XX P-P-SDB; AAU68525.

New polypeptides and nucleic acids, useful for diagnosis, treatment of  
PT inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone  
PT degenerative disorders, cancer and promoting wound healing  
XX  
XX Claim 1; Page 150-151; 336pp; English.

XX The invention relates to isolated human polypeptides (which may be  
CC cytokines) and the polynucleotides encoding them. The protein is useful  
CC for identifying a compound which binds to it (e.g. modulators, agonists  
CC and antagonists). The polynucleotides are useful as an array for mismatch  
CC detection. The proteins and nucleic acids are useful as nutritional  
CC sources or supplements. The protein exhibits activity relating  
CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,  
CC stem cell growth factor activity, immune stimulating or immune  
CC suppressing and activin or inhibin related activities. The proteins (and  
CC antibodies raised against them) and nucleic acids are therefore useful in  
CC the diagnosis and treatment of diseases and disorders such as cancer,  
CC central and peripheral nervous system diseases and neuropathies,  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular  
CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,

CC thrombocytopenia, stem cell disorders, aplastic anaemia, for  
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
CC growth, and in tissue repair, healing of burns, incisions, ulcers, for  
CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or  
CC periodontal disease, lung or liver fibrosis, reperfusion injury in  
CC various tissues, various immune deficiencies and disorders including  
CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,  
CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,  
CC such as asthma or other respiratory problems, coagulation disorders,  
CC haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory  
CC bowel disease, viral infection and are useful in altering bodily  
CC characteristics. The present sequence encodes a novel protein of the  
CC invention.

XX Sequence 745 BP; 157 A; 234 C; 235 G; 119 T; 0 other;

XX Query Match 51.3%; Score 321; DB 22; Length 745;

XX Best Local Similarity 100.0%; Pred. No. 4.8e-49;

XX Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 GGTGTGACGGGGGTGACGGGGGCGGATCTCTGAGCCATGGGGCGCGCG 170

DB 2 GGTGTGACGGGGGTGACGGGGGCGGATCTCTGAGCCATGGGGCGCGCG 61

QY 171 CGACGCGATCCGTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCT 230

DB 62 CGACGCGATCCGTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCT 121

QY 231 GAAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGCGGCGCATCCCGCGCGCGCTGCT 290

DB 122 GAAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGCGGCGCATCCCGCGCGCGCTGCT 181

QY 291 GTCATGAGACGCTTGACCTCAACGAGCTGCTGAGCTTCTACTGAGAGACTCAAG 350

DB 182 GTCATGAGACGCTTGACCTCAACGAGCTGCTGAGCTTCTACTGAGAGACTCAAG 241

QY 351 CGCGGAGCTCAACCGCTAACGTGCTGGCGGACATGGGCTTGCAGAGATGGCGGAGCT 410

DB 242 CGCGGAGCTCAACCGCTAACGTGCTGGCGGACATGGGCTTGCAGAGATGGCGGAGCT 301

QY 411 GCAGGCGGCAACGACACGAGG 431

DB 302 GCAGGCGGCAACGACACGAGG 322

RESULT 15 AAC98638 standard; cDNA; 432 BP.

AC AAC98638; XX

DE 09-MAR-2001 (first entry)

XX Human colon cancer antigen nucleotide sequence SEQ ID NO:648.

KW Human: colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytostatic; cardioactive; neuroprotective; vulnary;

KW immunomodulatory; muscular; gynaecological; gastrointestinal;

KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; renal disorder;

KW infectious disease; cardiovascular disorder; ss.

XX Homo sapiens.

XX MO200055351-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05883.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Ruben SM;  
XX WPI: 2000-587534/55.  
XX P-PSDB: AAB53881.

XX Colon cancer associated gene sequences, referred to as colon cancer  
XX antigens, useful for the treatment, prevention, and diagnosis of colon  
XX disorders such as colon cancer -  
XX Claim 1; Page 1175; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
XX called human colon cancer antigens, given in AAB5334 to AAB54006. The  
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;  
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
XX vulnerable, nephrotropic, antiinfective and antibacterial activities, and  
XX can be used in gene therapy. The colon cancer antigen polynucleotides,  
XX proteins and antibodies to the proteins are useful for the prevention,  
XX treatment and diagnosis of colon disorders, such as colon cancer. The  
XX polynucleotides may be used in diagnostics and research, such as for  
XX chromosome identification, and as hybridisation probes. The proteins  
XX may also be used to prevent diseases such as neural disorders, immune  
XX system disorders, muscular disorders, reproductive disorders, immune  
XX gastrointestinal disorders, wounds, renal disorders, infectious  
XX diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
XX AAB54007 represent sequences used in the exemplification of the present  
XX invention.

XX Sequence 432 BP; 77 A; 142 C; 152 G; 61 T; 0 other;

XX Query Match 51.1%; Score 320; DB 21; Length 432;

XX Best Local Similarity 100.0%; Pred. No. 7.4e-49;

XX Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 GCTGACGGGGGTGACGGGGGCGGATCTCTGAGCCATGGGGCGCGCGC 171

DB 8 GCTGACGGGGGTGACGGGGGCGGATCTCTGAGCCATGGGGCGCGCGC 67

QY 172 GACGCGATCCGTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCTG 231

DB 68 GACGCGATCCGTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCTG 127

QY 232 AAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGCGGCGCATCCCGCGGCGCGCTGCTG 291

DB 128 AAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGCGGCGCATCCCGCGGCGCGCTGCTG 187

QY 292 TCATGAGACGCTTGACCTCAACGAGCTGCTGAGCTTCTACTGAGAGACTCAAG 351

DB 188 TCATGAGACGCTTGACCTCAACGAGCTGCTGAGCTTCTACTGAGAGACTCAAG 247

QY 352 GCCGAGCTCAACCGCTAACGTGCTGGCGGACATGGGCTTGCAGAGATGGCGGAGCTG 411

DB 248 GCCGAGCTCAACCGCTAACGTGCTGGCGGACATGGGCTTGCAGAGATGGCGGAGCTG 307

QY 412 CAGGCGGCAACGACACGAGG 431

DB 308 CAGGCGGCAACGACACGAGG 327

Search completed: December 22, 2002, 05:27:28  
Job time : 263 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:24:11 ; Search time 52 seconds  
(without alignments)  
3691.916 Million cell updates/sec

Title: US-09-691-763B-4

Perfect score: 626  
Sequence: 1 agccctttgcctgcgaagc.....ggaagggaaggatcactt 626

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata1/1na/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata1/1na/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata1/1na/5A.COMB.seq:\*  
4: /cgn2\_6/ptodata1/1na/5B.COMB.seq:\*  
5: /cgn2\_6/ptodata1/1na/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata1/1na/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.6	8.4	44377	2	US-08-804-227C-7
2	52.6	8.4	44377	2	US-08-804-198-1
3	51.6	8.2	33529	4	US-09-144-085-3
4	51.2	8.2	2943	1	US-08-042-747A-7
5	50.8	8.1	30001	2	US-08-125-468-1
6	50.8	8.1	30001	2	US-08-474-933-1
7	50	8.0	2244	1	US-08-476-519-10
8	50	8.0	2244	5	PCT-US95-09323-10
9	50	8.0	2334	1	US-08-476-518-1
10	50	8.0	2334	5	PCT-US95-09323-1
11	49.4	7.9	1765	1	US-08-459-586-15
12	49.4	7.9	1765	2	US-08-282-696-15
13	49.4	7.9	18994	1	US-08-459-586-4
14	49.4	7.9	18994	2	US-08-282-696-4
15	48	7.7	530	3	US-08-758-662-4
16	48	7.7	2310	3	US-09-036-987A-25
17	48	7.7	2310	4	US-09-370-700-25
18	48	7.7	4524	2	US-08-843-998-7
19	48	7.7	4524	3	US-09-206-537-7
20	48	7.7	4524	2	US-09-430-854-7
21	47.4	7.6	800	5	PCT-US95-04801-4
22	47.4	7.6	50937	4	US-09-428-517-1
23	46.6	7.4	1656	4	US-09-385-028-14
24	46.6	7.4	1809	4	US-08-455-001-1
25	46.6	7.4	1809	5	PCT-US95-11869-1
26	46.6	7.4	15079	4	US-09-385-028-1
27	46.4	7.4	4188	4	US-07-751-891B-2

28	46.4	7.4	4242	3	US-07-705-490-2	Sequence 2, Appl
29	46.2	7.4	861	2	US-08-997-080-97	Sequence 97, Appl
30	46.2	7.4	861	2	US-08-997-362-97	Sequence 97, Appl
31	46.2	7.4	861	3	US-08-873-970-97	Sequence 97, Appl
32	46.2	7.4	861	4	US-09-095-855-97	Sequence 97, Appl
33	46.2	7.4	861	4	US-09-324-542-97	Sequence 97, Appl
34	46.2	7.4	861	4	US-09-205-426-97	Sequence 97, Appl
35	46	7.3	1506	4	US-09-158-767-7	Sequence 97, Appl
36	46	7.3	2261	4	US-09-158-767-1	Sequence 1, Appl
37	45.8	7.3	11219	1	US-07-642-734C-1	Sequence 1, Appl
38	45.8	7.3	11219	3	US-08-439-009A-1	Sequence 1, Appl
39	45.6	7.3	1227	4	US-09-385-028-23	Sequence 23, Appl
40	45.6	7.3	15079	4	US-09-385-028-1	Sequence 1, Appl
41	45.2	7.2	4692	2	US-08-916-917-1	Sequence 1, Appl
42	45.2	7.2	4692	2	US-08-972-631-1	Sequence 1, Appl
43	45.2	7.2	4692	2	US-08-972-629-1	Sequence 1, Appl
44	45.2	7.2	4692	2	US-08-972-630-1	Sequence 1, Appl
45	45.2	7.2	4692	2	US-08-672-211-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-804-227C-7  
; Sequence 7, Application US/08804227C  
; Patent No. 5876991  
; GENERAL INFORMATION:  
; APPLICANT: Dehoff, Bradley S.  
; APPLICANT: Kunstoss, Stuart A.  
; APPLICANT: Rostock, Paul R., Jr.  
; APPLICANT: Sutton, Kimberly L.  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THOMAS G. PLANT 1501  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY disk  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII(DOS) Text only  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,227C  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas, G.  
; REGISTRATION NUMBER: 35,784  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-2459  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 44377 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 350..14002  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 14046..20036  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 20110..31284  
; FEATURE:





DB 14170 AGCGGCTGCTGCACTGGCCGCGCTGCAAGAGCGCTGCTGCCGCGCAGACGGCC 14229  
QY 340 GAGACCTAGGGGCGCGCTACCGCTAACGCTGCTGCGCGAGAGGGCCCTGCGAGGAGT 399  
DB 14230 GAGCGCTCTCGCGGCTGCTGGCGCCGAGGTGAGAGGGGCGCTGCACTGCGAGAGCTG 14289  
QY 400 GCGGCGCAGCTGCA 413  
DB 14290 ACGCGGAGCTGGA 14303

RESULT 4  
US-08-042-747A-7  
Sequence 7, Application US/08042747A  
Patent No. 5487969  
GENERAL INFORMATION:  
APPLICANT: Eberle, Richard  
APPLICANT: Black, Darla  
APPLICANT: Schlickefiel, Franco  
APPLICANT: Hilliard, Julia K.  
TITLE OF INVENTION: Cloning and Amplification of Monkey B  
NUMBER OF INVENTIONS: 12  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cox & Smith Incorporated  
STREET: 112 East Pecan Street, Suite 2000  
CITY: San Antonio  
STATE: Texas  
COUNTRY: USA  
ZIP: 78205

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/042,747A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Haymond, W. Bradley  
REGISTRATION NUMBER: 35186  
REFERENCE/DOCKET NUMBER: S-0072.179  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 210-554-5500  
TELEFAX: 210-226-8395  
TELEX: 767609

INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2943 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 87..2744  
US-08-042-747A-7

Query Match 8.2%; Score 51.2; DB 1: Length 2943;  
Best Local Similarity 44.5%; Pred. No. 0.064;  
Matches 203; Conservative 0; Mismatches 253; Indels 0; Gaps 0;

QY 23 CGGACCGGGGCGGAGTCCGGGAGACCAAGTGGAGGAGGAGCGGGAGTCCAGGTTCCG 82  
DB 208 CCGCCACGCGCGCTTCCGAGAGAGAGGTCGCCGACGAGGAGGAGGAGGAGGCTG 267  
QY 83 CCGCGAGCGCGACTTCTCTGCTGCTGCGCGCTGCGAGCGGCGGAGCGCGCAAGCGCG 142  
DB 268 CCAACCG 327  
QY 143 GGGATCTTGGAGCCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 202

DB 328 ACTGCGGAGATCAAGGCCCGGGAGCGAGCGCACTTACGCTGTGCCCGCGCCGA 387  
QY 203 CGCGGAGGAGCTCAAGATCTCAAGCTGCAAGGCTGCTGCGCGCTCGCGCGAGGCGT 262  
DB 368 CCGGCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447  
QY 263 ACGGCGCATCCCGCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 322  
DB 448 AGACTACAGGAGGAGATCGCGCTGCTTCAAGGAGAACATCGCCCGCTACAGATTCA 507  
QY 323 TGGTACGCTTCTACCTGAGACCTACGCGCGCGCGCGCGCGCTAACGCTGCTGCGCGCA 382  
DB 508 AGGCGCACATGTACTACAGAGAGCTGACCGTCTGCGAGGTCTGCTGCTGCTGCTGCTGCT 567  
QY 383 TGGGCTGCAAGATGCGCGCGAGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 442  
DB 568 CGCAGTTTCATGGGATCTTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 627  
QY 443 CGTCCCGCTCAACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478  
DB 628 AGATCAAGCCAAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663

RESULT 5  
US-08-125-468-1/c  
Sequence 1, Application US/08125468  
Patent No. 5589385  
GENERAL INFORMATION:  
APPLICANT: Ryan, Michael J.  
APPLICANT: Lovlin, Jason A.  
APPLICANT: Strathy, Nancy  
APPLICANT: Fanti, Susan E.  
TITLE OF INVENTION: Cloning of the biosynthetic pathway for  
TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmid  
NUMBER OF INVENTIONS: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07470

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125,468  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Iseydos, Estelle J.  
REGISTRATION NUMBER: 31,145  
REFERENCE/DOCKET NUMBER: 31,255-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201)831-3241  
TELEFAX: (201)831-3305  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30001 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-125-468-1

Query Match 8.1%; Score 50.8; DB 1: Length 30001;  
Best Local Similarity 45.6%; Pred. No. 0.083;  
Matches 219; Conservative 0; Mismatches 257; Indels 4; Gaps 1;

QY 3 CGCCTTTTCTGAGAGGCAACGAGCGGCGCGGAGCTCGGAGACCAAGTGGAGAGAA 62

```
Db 24497 GGGGGGCTTCCAGAGGCGCATGTGCTCGCCGAGCGGAGTGGAGCCGGGGATGACCGAC 24438
Oy 63 GGGGGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGTGCGCGCTGACGCGG 122
Db 24437 GTCTGTGTGACCGGATCCGC-----GTGTGTGACCGAACGAGCTGGGGCTGGGAGTA 24382
Oy 123 GTGAGCGGCGGAGCGCGGGGATCTGAGAGCCATGGGGCGCGCGGAGCGGATCCT 182
Db 24381 CTGGGAGCGCACACTGGCGGGGCTCGGTCTGGGCGCTGACCCGCTGTGACTCCAC 24322
Oy 183 GGATGCGCTGAGAGACCTGACCGCGGAGAGCTTCAGAACTTCAAGCTGAAGCTGCTC 242
Db 24321 CGGGTACCGGGGCGCACGTGCGCGGAGGTGACCTTCAACCGAGGAACACCTGCGGG 24262
Oy 243 GGTGCGCTGCGGAGGAGGTACGGGCGCATCCGCGGGGCGGCTCTGTGCTGATGAGCG 302
Db 24261 CCGACTGCTCCCGACACGACGACACATGACGCGGCTGCGCTGTGCGCGCGGAGAGC 24202
Oy 303 CTGAGACTTACCGACACAGCTGTGCTTCTTACCTGAGAGACCTTACGCGCGGAGCTTAC 362
Db 24201 GCTGCGCGAGCGCGGGGCGGACCCGACGACGATGCCGACTTCTCGCGGCTGTGAC 24142
Oy 363 CGTAACTGCTCGGACATGAGCTGCTGACGAGATGGCGGAGCTGACGCGCGGAC 422
Db 24141 GCGCGCTCGCGCGCGGCGGCTTGAAGTTGCGCGAGAGAACTGACGCGCTGTGAGCAA 24082
Oy 423 GCACGAGGTGAGCGCGCGCGCTCCCTCCACCCGCTCTTCCCTCCACACACACAG 482
Db 24081 GGGCGCGGCTTACGTACGTCAGGCTTACGACTCTTCCGCTGTCTACCGCTGACACCGG 24022

RESULT 6
US-08-474-933-1/c
; Sequence 1, Application US/08474933
; Patent No. 586410
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fanti, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmids
; TITLE OF INVENTION: useful therein
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,933
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,468
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Teevados, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3241
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS: 1
```

```
; LENGTH: 30001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-474-933-1
Query Match 8.1%; Score 50.8; DB 2; Length 30001;
Best Local Similarity 45.6%; Pred. No. 0.083;
Matches 219; Conservative 0; Mismatches 257; Indels 4; Gaps 1;

Oy 3 CGCTTTTCTGAGAGCGACGACGAGCGGGGAGTGGGAGACCAAGTGGAGGAA 62
Db 24497 GGGGGGCTTCCAGAGGCGCATGTGCTCGCCGAGCGGAGTGGAGCCGGGGATGACCGAC 24438
Oy 63 GGGGGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGTGCGCGCTGACGCGG 122
Db 24437 GTCTGTGTGACCGGATCCGC-----GTGTGTGACCGAACGAGCTGGGGCTGGGAGTA 24382
Oy 123 GTGAGCGGCGGAGAGCTGACCGCGGAGAGCTTCAGAACTTCAAGCTGAAGCTGCTC 182
Db 24321 CGGGTACCGGGGCGCACGTGCGCGGAGGTGACCTTCAACCGAGGAACACCTGCGGG 24262
Oy 183 GGATGCGCTGAGAGACCTGACCGCGGAGAGCTTCAGAACTTCAAGCTGAAGCTGCTC 242
Db 24321 CCGACTGCTCCCGACACGACGACACATGACGCGGCTGCGCTGTGCGCGCGGAGAGC 24202
Oy 243 GGTGCGCTGCGGAGGAGGTACGGGCGCATCCGCGGGGCGGCTCTGTGCTGATGAGCG 302
Db 24261 CCGACTGCTCCCGACACGACGACACATGACGCGGCTGCGCTGTGCGCGCGGAGAGC 24202
Oy 303 CTGAGACTTACCGACACAGCTGTGCTTCTTACCTGAGAGACCTTACGCGCGGAGCTTAC 362
Db 24201 GCTGCGCGAGCGCGGGGCGGACCCGACGACGATGCCGACTTCTCGCGGCTGTGAC 24142
Oy 363 CGTAACTGCTCGGACATGAGCTGCTGACGAGATGGCGGAGCTGACGCGCGGAC 422
Db 24141 GCGCGCTCGCGCGCGGCGGCTTGAAGTTGCGCGAGAGAACTGACGCGCTGTGAGCAA 24082
Oy 423 GCACGAGGTGAGCGCGCGCGCTCCCTCCACCCGCTCTTCCCTCCACACACACAG 482
Db 24081 GGGCGCGGCTTACGTACGTCAGGCTTACGACTCTTCCGCTGTCTACCGCTGACACCGG 24022

RESULT 7
US-08-476-519-10
; Sequence 10, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 5750876e1 Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BBAF
; STREET: 700 Chesterfield Parkway No. 5750876th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
```

```

;
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
;
US-08-476-519-10

```

```

Query Match      8.0%; Score 50; DB 1; Length 2244;
Best Local Similarity 44.9%; Pred. No. 0.11;
Matches 230; Conservative 0; Mismatches 280; Indels 2; Gaps 1;

```

```

QY 15 GAGGGCAACGACCGGGGAGTCGGGAGACCAAGTGGAGGAAGCGGGAGTCC 74
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1339 GAGTGGAAAGCGCTCTACCGCGACCGCGTCGCCAAGAAAGCAAGCTCGCGTGAA 1398
QY 75 AGGTTCGCCCCGGAGCCGACCTTCCTGTCGTCGCGCCGCTCAGCGGGGTAGCGCGGC 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1399 ACGGTACACCCCGGACCGCTGCGCACGCGCTTCGCGGCTCCAAAGACGTCGTCGCGAC 1458
QY 135 AGCGCGCGGGGATCTGCG--AGCCATGGGGCGCGCGCGACGCGCATCTGATGCGCTG 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1459 GACGGCGCGAAGCGCTGCGCATCTGATCACTTCGTCGTCGCCACGACGCGCTTACCCCTC 1518
QY 193 GAGAACTGACCGCGCGAGAGCTCAAGAAATTCAAGCTGAAGCTGCTGCGTCCGCTG 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1519 AAGCACTGTACGCTTCAACGACAAAGCAAGCAACGACCGCTGCGCGCTGCGCTGCTC 1578
QY 253 CGCAGAGGCTACGAGCGCATCCCGCGGGGCGCGCTGCTCATGAGACGCTTGGACCTC 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1579 GACGGCGCGAGGAGCAACAACCTGAGCTGGAACAGGGCGGCTGTCGCGCGAGCGCG 1638
QY 313 ACCGACAAGCTGTGCTTACCTTCTAGAGACCTACGCGCGCGACGCTACCGCTAACGTCG 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1639 AAGGCGCGGACCGGACGCTGCTGATGCTCAAGCGCGCGCGCTGCGGATGATCAC 1698
QY 373 CTGCGGACATGGGCTTGCAGAGATGGCGGAGCTGCGAGGCGCGCACAGAGGT 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1699 GCGGCGGACGAGGCGCTGCGACAGTTCCGGCAACAACAACCTCAACCTGGAATTG 1758
QY 433 GAGCGCGCCCGCTTCCCTCCACCCGCTCTTCCCTCCACCCACAGAGCTTACCC 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1759 GCGGCAACGTGGCTGTACTGTAGAGCGGAGCGGCTGAGAGCGACAGACCTACACC 1818
QY 493 GCGGGCTCTTCCGCTTCTGTTCTCTCTACCC 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1819 AAGCGCTGATGCGCTTCCGCAAGGCGACACC 1850

```

```

RESULT 8
PCT-US95-09323-10
; Sequence 10, Application PC/TUS9509323
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

;
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09323
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
;
PCT-US95-09323-10

```

```

Query Match      8.0%; Score 50; DB 5; Length 2244;
Best Local Similarity 44.9%; Pred. No. 0.11;
Matches 230; Conservative 0; Mismatches 280; Indels 2; Gaps 1;

```

```

QY 15 GAGGGCAACGACCGGGGAGTCGGGAGACCAAGTGGAGGAAGCGGGAGTCC 74
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1339 GAGTGGAAAGCGCTCTACCGCGACCGCGTCGCCAAGAAAGCAAGCTCGCGTGAA 1398
QY 75 AGGTTCGCCCCGGAGCCGACCTTCCTGTCGTCGCGCCGCTCAGCGGGGTAGCGCGGC 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1399 ACGGTACACCCCGGACCGCTGCGCACGCGCTTCGCGGCTCCAAAGACGTCGTCGCGAC 1458
QY 135 AGCGCGCGGGGATCTGCG--AGCCATGGGGCGCGCGCGACGCGCATCTGATGCGCTG 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1459 GACGGCGCGAAGCGCTGCGCATCTGATCACTTCGTCGTCGCCACGACGCGCTTACCCCTC 1518
QY 193 GAGAACTGACCGCGCGAGAGCTCAAGAAATTCAAGCTGAAGCTGCTGCGTCCGCTG 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1519 AAGCACTGTACGCTTCAACGACAAAGCAAGCAACGACCGCTGCGCGCTGCGCTGCTC 1578
QY 253 CGCAGAGGCTACGAGCGCATCCCGCGGGGCGCGCTGCTCATGAGACGCTTGGACCTC 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1579 GACGGCGCGAGGAGCAACAACCTGAGCTGGAACAGGGCGGCTGTCGCGAGAGCGCG 1638
QY 313 ACCGACAAGCTGTGCTTACCTTCTAGAGACCTACGCGCGCGACGCTACCGCTAACGTCG 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1639 AAGGCGCGGACCGGACGCTGCTGATGCTCAAGCGCGCGCGCTGCGGATGATCAC 1698
QY 373 CTGCGGACATGGGCTTGCAGAGATGGCGGAGCTGCGAGGCGCGCACAGAGGT 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1699 GCGGCGGACGAGGCGCTGCGACAGTTCCGGCAACAACAACCTCAACCTGGAATTG 1758
QY 433 GAGCGCGCCCGCTTCCCTCCACCCGCTCTTCCCTCCACCCACAGAGCTTACCC 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1759 GCGGCAACGTGGCTGTACTGTAGAGCGGAGCGGCTGAGAGCGACAGACCTACACC 1818
QY 493 GCGGGCTCTTCCGCTTCTGTTCTCTCTACCC 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1819 AAGCGCTGATGCGCTTCCGCAAGGCGACACC 1850

```

```

RESULT 9
US-08-476-519-1
; Sequence 1, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 3750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

```

```
ADDRESS: Grace L. Bonner, Monsanto Company, B44F
STREET: 700 Chesterfield Parkway No. 5750876th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,519
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/281902
FILING DATE: 28-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(13577)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2334 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-476-519-1
```

Query Match 8.0%; Score 50; DB 1; Length 2334;

Best Local Similarity 44.9%; Pred. No. 0.11; Matches 230; Conservative 0; Mismatches 280; Indels 2; Gaps 1;

```

15 GAGGCAACGACCGGCGGGAGTGGGAGACCAAGTGGAGGAGGCGGGAGTCC 74
1432 GAGTGGAAAGCGCTTACCGGACGCGCTCGGCAAGAACGAAAGCTCGCGTGA 1491
75 AGGTTCCGCGCGGAGCGGACCTCTGTTGGGGGCTGAGAGGGGTTGAGGGG 134
1492 ACGGTACCGCGGCGGACCTGAGCCACGCGCTTCCAGGACCTGTACGGGAGC 1551
135 AGCGCGCGGAGTCTGCG--AGCATGGGGCGCGCGGACGCGCATCTGATGCGCTG 192
1552 GACGGCGGAAAGCGGCTGATCGATCACTGCTGATGCTGCGCCACGAGCGCTTCA 1611
193 GAGAACCTGACCGCGGAGAGCTCAAGAGTGAAGCTGAAGCTGTGCTGCGGCTG 252
1612 AAGCACTGTACCGCTACACAGACACAGACACAGCCGTCGATGAGGGCGCTCC 1671
253 CGGAGGGCTACGGGCGGATCCCGGGGCGCGCTGCTCATGAGACCTTGGACCTC 312
1672 GACGGCGGAGAGCGACCAACCTGAGCTGGAAACAGAGCGGCGATGCTGCGAC 1731
313 ACCGACAAGCTGATGAGCTTCTACCTGAGACCTACGGGCGGAGCTCAACGCTG 372
1732 AAGCGCGCGGCGGACCGGACTGCGCTGTGCTGATGCTGACGGCGGCTGCTGAT 1791
373 CTGCGGACATGAGCTGACAGAAATGAGCCGGGAGCTGACAGGCGGCGACACGAG 432
1792 GCGGCGGAGAGAGCGCTGCGACCGAGTTCGGGCAACACAACTTACAACTGATTCG 1851
433 GAGCGCGCGCGCTTCCCGTCCACCGCTTTTCCCTCCACACAGAGGCTTACGCC 492
1852 GCGGCAACTGCTGTACTGAGAGCGGACGCGGCTGAGAGCGGACGACGACTTAC 1911
493 GCGGCTCTTCCGCTTCTGTTCTCTTCTACCC 524
1912 AAGCGCTGATGCGCTTCCGCAAGGCGCACCC 1943
```

```

RESULT 10
PCT-US95-09323-1
Sequence 1, Application PC/7US9509323
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
TITLE OF INVENTION: Containing It and Methods of using Isoamylases
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09323
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/281902
FILING DATE: 28-JUL-1994
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2334 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-09323-1
```

Query Match 8.0%; Score 50; DB 5; Length 2334;

Best Local Similarity 44.9%; Pred. No. 0.11; Matches 230; Conservative 0; Mismatches 280; Indels 2; Gaps 1;

```

15 GAGGCAACGACCGGCGGGAGTGGGAGACCAAGTGGAGGAGGCGGGAGTCC 74
1432 GAGTGGAAAGCGCTTACCGGACGCGCTCGGCAAGAACGAAAGCTCGCGTGA 1491
75 AGGTTCCGCGCGGAGCGGACCTCTGTTGGGGGCTGAGAGGGGTTGAGGGG 134
1492 ACGGTACCGCGGCGGACCTGAGCCACGCGCTTCCAGGACCTGTACGGGAGC 1551
135 AGCGCGCGGAGTCTGCG--AGCATGGGGCGCGCGGACGCGCATCTGATGCGCTG 192
1552 GACGGCGGAAAGCGGCTGATCGATCACTGCTGATGCTGCGCCACGAGCGCTTCA 1611
193 GAGAACCTGACCGCGGAGAGCTCAAGAAATTCAGAGCTGAAGCTGTGCTGCGGCTG 252
1612 AAGCACTGTACCGCTACACAGACACAGACACAGCCGTCGATGAGGGCGCTCC 1671
253 CGGAGGGCTACGGGCGGATCCCGGGGCGCGCTGCTCATGAGACCTTGGACCTC 312
1672 GACGGCGGAGAGCGACCAACCTGAGCTGGAAACAGAGCGGCGATGCTGCGGAG 1731
313 ACCGACAAGCTGATGAGCTTCTACCTGAGACCTACGGGCGGAGCTCAACGCTG 372
1732 AAGCGCGCGGCGGACCGGACTGCGCTGTGCTGATGCTGACGGCGGCTGCTGAT 1791
373 CTGCGGACATGAGCTGACAGAAATGAGCCGGGAGCTGACAGGCGGCGACACGAG 432
1792 GCGGCGGAGAGAGCGCTGCGACCGAGTTCGGGCAACAACTTACAACTGATTCG 1851
433 GAGCGCGCGCGCTTCCCGTCCACCGCTTTTCCCTCCACACAGAGGCTTACGCC 492
1852 GCGGCAACTGCTGTACTGAGAGCGGACGCGGCTGAGAGCGGACGACGACTTAC 1911
493 GCGGCTCTTCCGCTTCTGTTCTCTTCTACCC 524
1912 AAGCGCTGATGCGCTTCCGCAAGGCGCACCC 1943
```

RESULT 11





Db 14384 TGTGCACTTTCTGCGCGGCGCTGCTG9GCGCGCTG9GCGTGGAGCGCAAGTCGCGC 14443  
QY 323 TGTGCACTTTCTGAGACCTAGCGGCGGAGTCAAGCTACGCTGCGGAGACA 382  
Db 14444 CCGCGCGCGGAGCGCGCGGCGGCTGCTGCGCGAGCTGGAGCGCGAGCAATGGAGCGGC 14503  
QY 383 TGGGCTCTGAGAGATGCGCGGCGAGCTGCGAGCGCGCAAGCGAGGTGAGCGCGCC 441  
Db 14504 CGCGTGGAGACGCGTGCACCTGCACCGCGCGCGCGCTGTGGCGCGGAGCGCGCAC 14562

## RESULT 15

US-08-758-662-4  
/ Sequence 4, Application US/08758662  
/ Patent No. 6114150  
/ GENERAL INFORMATION:  
/ APPLICANT: Weisman, Sherman  
/ APPLICANT: Baskaran, Namadev  
/ TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACIDS  
/ NUMBER OF SEQUENCES: 8  
/ CORRESPONDENCE ADDRESSES:  
/ ADDRESSEE: SEED and BERRY LLP  
/ STREET: 701 Fifth Avenue, 6300 Columbia Center  
/ CITY: Seattle  
/ STATE: Washington  
/ COUNTRY: USA  
/ ZIP: 98104  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patentln Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/758,662  
/ FILING DATE: 29-NOV-1996  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: No. 6114150tendurg Ph.D., Carol  
/ REGISTRATION NUMBER: 39,317  
/ REFERENCE/DOCKET NUMBER: 390036.402C1  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (206) 622-4900  
/ TELEFAX: (206) 682-6031  
/ INFORMATION FOR SEQ ID NO: 4:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 330 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ US-08-758-662-4

Query Match 7.7%; Score 48; DB 3; Length 530;  
Best Local Similarity 55.4%; Pred. No. 0.25; Mismatches 75; Indels 0; Gaps 0;  
Matches 93; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 23 CGGACGCGGCGCGGAGTCGCGGAGACCAAGTGGAGAGCGGCGGAGTCCAGGTTCCG 82  
Db 310 CGGCGCGGCGCGGCGG 369  
QY 83 CCGCGGAGCGGAGTCT 142  
Db 370 CCGCGGCGCGGCG 429  
QY 143 GGGATCTTGGAGCATGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 190  
Db 430 GGGGCGCGGCG 477

Search completed: December 22, 2002, 06:48:32  
Job time : 142 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:23:21 / Search time 1992 seconds  
(without alignments)  
5089.546 Million cell updates/sec

Title: US-09-691-763b-4

Perfect score: 626  
Sequence: 1 agcgcttctgtcgtgaagc.....ggaagggaaggatcactt 626

Scoring table: IDENTITY\_NUC  
Gapop 10.0 / Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmuj:\*  
5: em\_estov:\*  
6: em\_estopl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_iny:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	424.4	67.8	866 12	BG684386
2	340.4	54.4	741 13	BM051141
3	338	54.0	767 14	BM910918
4	337.2	53.9	1079 13	BM456838
5	321	51.3	744 12	BG255521
6	320	51.1	320 13	BT756931

7	320	51.1	930 13	BM459678
8	317	50.6	830 12	BF238010
9	316	50.5	792 13	BM549665
10	315	50.3	804 14	BM926438
11	314	50.2	750 12	BG764161
12	314	50.2	856 12	BG767422
13	312.4	49.9	743 12	BG337806
14	301.2	48.1	1294 14	BM918855
15	296	47.3	446 14	BM726665
16	270.4	43.2	634 14	AT991236
17	270	43.1	687 12	BE909218
18	259	41.4	888 12	BE906286
19	258.4	41.3	678 9	AT148558
20	248	39.6	676 14	BM972785
21	247	39.5	675 12	BE908204
22	244.4	39.0	615 9	AT146818
23	241.8	38.6	267 13	BT025158
24	226.6	36.2	308 14	H16108
25	224.4	35.8	262 14	W17108
26	222.8	35.6	560 13	BM364268
27	221.4	35.4	1093 14	BM920038
28	220.4	35.2	830 10	BE560228
29	218.6	34.9	723 10	BE560247
30	218.4	34.9	555 10	AV664060
31	214	34.2	446 9	AA442018
32	211.6	33.8	506 13	BM105270
33	211.2	33.7	404 10	AM479270
34	206	32.9	448 13	BT339834
35	206	32.9	633 14	BM681310
36	205.2	32.8	638 14	BM988354
37	204	32.6	417 13	BM363061
38	204	32.6	427 13	BM365021
39	204	32.6	469 13	BM363361
40	204	32.6	565 13	BM362328
41	200.4	32.0	352 13	BM365886
42	200.4	32.0	367 13	BM365745
43	200.4	32.0	412 13	BM365704
44	200.4	32.0	456 13	BM365720
45	200.4	32.0	473 13	BM255820

#### ALIGNMENTS

RESULT 1  
LOCUS BG684386 866 bp mRNA linear EST 01-MAY-2001  
DEFINITION 602636137F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4763985 5',  
mRNA sequence.  
ACCESSION BG684386  
VERSION BG684386.1 GI:13915783  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 866)  
AUTHORS NIH-MGC  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNLT Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs@rmail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: LLCM1620 row: i column: 10  
High quality sequence stop: 862.  
Location/Qualifiers  
L. 866

FEATURES  
Source

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4763985"  
 /clone\_lib="NIH\_MGC\_48"  
 /tissue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: B-cells; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(6). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

BASE COUNT 154 a 279 c 310 g 123 t  
 ORIGIN

Query Match 67.8%; Score 424.4; DB 12; Length 866;  
 Best Local Similarity 99.8%; Pred. No. 3.8e-69;  
 Matches 425; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 CTTTTCCTGAGGCGCAAGGACCGGGGCGGAGTCTGGAGACCAAGTGGAGGAGGAGGC 65  
 DB 132 CTTTTCCTGAGGCGCAAGGACCGGGGCGGAGTCTGGAGACCAAGTGGAGGAGGAGGC 211  
 OY 66 GGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGCTCGCGCGCTGCAGCGGGGTG 125  
 DB 212 GGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGCTCGCGCGCTGCAGCGGGGTG 271  
 OY 126 AGCGGGGCGAGCGGGGCGGAGTCTGGAGACCAAGTGGAGGAGGAGGAGGC 185  
 DB 272 AGCGGGGCGAGCGGGGCGGAGTCTGGAGACCAAGTGGAGGAGGAGGAGGC 331  
 OY 186 TGGCTCTGAGAACCTTACCGCCGAGAGCTTCAGCTTGAAGCTGCTGTGCTGT 245  
 DB 332 TGGCTCTGAGAACCTTACCGCCGAGAGCTTCAGCTTGAAGCTGCTGTGCTGT 391  
 OY 246 GCCCTCTGAGAACCTTACCGCCGAGAGCTTCAGCTTGAAGCTGCTGTGCTGT 305  
 DB 392 GCCCTCTGAGAACCTTACCGCCGAGAGCTTCAGCTTGAAGCTGCTGTGCTGT 451  
 OY 306 GGACCTACCGAGAACCTTGAAGCTTCTACCTTGGAGACCTTACGGGCGCCAGCTCACCGC 365  
 DB 452 GGACCTACCGAGAACCTTGAAGCTTCTACCTTGGAGACCTTACGGGCGCCAGCTCACCGC 511  
 OY 366 TTAAGTCTGCGGAGACATGGCTCTGAGAGATGGCCGGGACATCTGAGGCGCCAGCA 425  
 DB 512 TTAAGTCTGCGGAGACATGGCTCTGAGAGATGGCCGGGACATCTGAGGCGCCAGCA 571  
 OY 426 CCAGG 431  
 DB 572 CCAGG 577

RESULT 2  
 LOCUS BM051141 741 bp mRNA linear EST 07-NOV-2001  
 DEFINITION 603634063F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:5424590 5',  
 mRNA sequence.  
 ACCESSION BM051141  
 VERSION BM051141.1 GI:16780408  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 741)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: InCyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LNCM185 row: n column: 15  
 High quality sequence stop: 643.  
 Location/Qualifiers

FEATURES  
 source

1. 741  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5424590"  
 /clone\_lib="NIH\_MGC\_43"  
 /tissue\_type="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(5). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

BASE COUNT 139 a 243 c 237 g 122 t

Query Match 54.4%; Score 340.4; DB 13; Length 741;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-53;  
 Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 90 GCCGACTTCCTCTGCTGCGGCGGCTGACAGCGGCGGCGGCGGAGCC 149  
 DB 1 GCCGACTTCCTCTGCTGCGGCGGCTGACAGCGGCGGCGGAGCC 60  
 OY 150 TGGAGCATGGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 209  
 DB 61 TGGAGCATGGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
 OY 210 GGAGGCAAGAGTTCAAGCTGAAGCTGCTGCGGCGGCGGCGGCGGCGG 269  
 DB 121 GAGCTCAAGAGTTCAAGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 180  
 OY 270 CATCCCGGGGGCGGCTGTCTCATGAGCCCTTGGACCTCACGAGAGCTGTAG 329  
 DB 181 CATCCCGGGGGCGGCTGTCTCATGAGCCCTTGGACCTCACGAGAGCTGTAG 240  
 OY 330 CTTTACCTGGAGACCTTACGCGCGGAGCTCAACGCTGAGCTGGAGATGGGCT 389  
 DB 241 CTTTACCTGGAGACCTTACGCGCGGAGCTCAACGCTGAGCTGGAGATGGGCT 300  
 OY 390 GCAGGAGATGGCGGCGGAGCTGAGGCGGCGGCGGCGGCGGCGGCGG 431  
 DB 301 GCAGGAGATGGCGGCGGAGCTGAGGCGGCGGCGGCGGCGGCGGCGG 342

RESULT 3  
 LOCUS BM910918 767 bp mRNA linear EST 12-MAR-2002  
 DEFINITION AGENCOURT\_6615922 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5454497  
 5', mRNA sequence.  
 ACCESSION BM910918  
 VERSION BM910918.1 GI:19361297  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 767)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution by: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM1949 row: 1 column: 18  
High quality sequence stop: 641.

## FEATURES

## source

Location/Qualifiers

1..767  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5454497"  
/clone\_1lb="NIH\_MGC\_98"  
/tissue\_type="astrocytoma grade IV, cell line"  
/note="Organ: brain; Vector: pOT87; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5'-adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 163 a 241 c 237 g 124 t 2 others  
ORIGIN

Query Match 54.0%; Score 338; DB 14; Length 767;  
Best Local Similarity 100.0%; Pred. No. 3.6e-53;  
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 ACTTCTCTCTGTCGCGGCTGCAGCGGGGTGAGCGCGCGAGCGCGCGGATCTTGGA 153  
Db 1 ACTTCTCTCTGTCGCGGCTGCAGCGGGGTGAGCGCGCGAGCGCGCGGATCTTGGA 60  
QY 154 GCCATGGGGCGCGCGCGCGGACGCCATCTTGATGCGCTGAGAACTGACCGCGAGAG 213  
Db 61 GCCATGGGGCGCGCGCGCGGACGCCATCTTGATGCGCTGAGAACTGACCGCGAGAG 120  
QY 214 CTCAGAGATTCAAGCTGAGCTGCTGTGCGCGCGCTGCGCGAGGGCTACGGGGCATTC 273  
Db 121 CTCAGAGATTCAAGCTGAGCTGCTGTGCGCGCGCTGCGCGAGGGCTACGGGGCATTC 180  
QY 274 CCGCGGGCGCGCTGCTGTCCATGAGCGCTTGACCTACCGAAGCTGTGTCAGCTTC 333  
Db 181 CCGCGGGCGCGCTGCTGTCCATGAGCGCTTGACCTACCGAAGCTGTGTCAGCTTC 240  
QY 334 TACCTGAGACCTAGCGCGCGGAGCTACCGCTACGCTGCGCGAGCATGGGCTGCAG 393  
Db 241 TACCTGAGACCTAGCGCGCGGAGCTACCGCTACGCTGCGCGAGCATGGGCTGCAG 300  
QY 394 GAGATGGCGGGGAGCTGACAGGGCGCCACGACCCAGG 431  
Db 301 GAGATGGCGGGGAGCTGACAGGGCGCCACGACCCAGG 338

RESULT 4  
BMA56838 1079 bp mRNA linear EST 05-FEB-2002  
LOCUS BMA56838  
DEFINITION AGENCOURT\_6404153 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5583760  
5', mRNA sequence.  
ACCESSION BMA56838  
VERSION BMA56838.1 GI:18505878  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1079)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE  
JOURNML  
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution by: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM12347 row: f column: 17  
High quality sequence stop: 682.

## FEATURES

## source

Location/Qualifiers

1..1079  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5583760"  
/clone\_1lb="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."  
BASE COUNT 246 a 360 c 297 g 171 t 5 others  
ORIGIN

Query Match 53.9%; Score 337.2; DB 13; Length 1079;  
Best Local Similarity 99.1%; Pred. No. 4.9e-53;  
Matches 339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 95 CTTCCTCTCTGTCGCGGCTGCAGCGGGGTGAGCGCGCGAGCGCGCGGATCTTGAG 154  
Db 1 CTTCCTCTCTGTCGCGGCTGCAGCGGGGTGAGCGCGCGAGCGCGCGGATCTTGAG 60  
QY 155 CCATGGGGCGCGCGCGCGGACGCCATCTTGATGCGCTGAGAACTGACCGCGAGAGG 214  
Db 61 CCATGGGGCGCGCGCGCGGACGCCATCTTGATGCGCTGAGAACTGACCGCGAGAGG 120  
QY 215 TCAAGAGATTCAAGCTGAGCTGCTGTGCGCGCGCTGCGCGAGGGCTACGGGGCATTC 274  
Db 121 TCAAGAGATTCAAGCTGAGCTGCTGTGCGCGCGCTGCGCGAGGGCTACGGGGCATTC 180  
QY 275 CCGCGGGCGCGCTGCTGTCCATGAGCGCTTGACCTACCGAAGCTGTGTCAGCTTC 334  
Db 181 CCGCGGGCGCGCTGCTGTCCATGAGCGCGCTTGACCTACCGAAGCTGTGTCAGCTTC 240  
QY 335 ACCTGAGACCTAGCGCGCGGAGCTACCGCTACGCTGCGCGAGCATGGGCTGCAGG 394  
Db 241 ACCTGAGACCTAGCGCGCGGAGCTACCGCTACGCTGCGCGAGCATGGGCTGCAGG 300  
QY 395 AGATGGCGGGGAGCTGACAGGGCGCCACGACCGGTTAGC 436  
Db 301 AGATGGCGGGGAGCTGACAGGGCGCCACGACCGGCTTCG 342

RESULT 5  
BG255521 744 bp mRNA linear EST 13-FEB-2001  
LOCUS BG255521  
DEFINITION 602367671F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4476088 5',  
mRNA sequence.  
ACCESSION BG255521  
VERSION BG255521.1 GI:12765259  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
 COMMENT  
 Unpublished (1999)  
 Contact: Robert Straussberg, Ph.D.  
 Email: rgsaps-remail.nih.gov  
 Tissue Procurement: DCTD/OTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM10302 row: m column: 17  
 High quality sequence stop: 66L.  
 Location/Qualifiers  
 source  
 1..744  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4476088"  
 /clone\_1lb="NIH\_MGC\_91"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: prostate; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; Oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."  
 BASE COUNT  
 159 a 227 c 238 g 120 t  
 ORIGIN  
 Query Match 51.3% Score 321: DB 12: Length 744:  
 Best Local Similarity 100.0%, Pred. No.5,1e-50;  
 Matches 321: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
 OY 111 GGCTGCACGGGGGTGAGCGGCGGAGCGCGCGGGAGACTCTGAGACCATGAGGCGCGCGC 170  
 Db 1 GGCTGCACGGGGGTGAGCGGCGGAGCGCGCGGGAGACTCTGAGACCATGAGGCGCGCGC 60  
 OY 171 CGAGCGCACTCTGGATGCGCTGAGAACCTGACCGCGGAGAGACTCAAGAGTTCACAGCT 230  
 Db 61 CGAGCGCACTCTGGATGCGCTGAGAACCTGACCGCGGAGAGACTCAAGAGTTCACAGCT 120  
 OY 231 GAGCTGCTGTGCGGTGCGCGCGAGGAGGCTACGGGCGGAGCTCCGCGGCGCGCTGCT 290  
 Db 121 GAGCTGCTGTGCGGTGCGCGCGAGGAGGCTACGGGCGGAGCTCCGCGGCGCGCTGCT 180  
 OY 291 GTCATGAGAGCCCTTGGACCTCAGCAGCAACTGCTGTCAGCTTCTACCTGAGAGACTTACAGG 350  
 Db 181 GTCATGAGAGCCCTTGGACCTCAGCAGCAACTGCTGTCAGCTTCTACCTGAGAGACTTACAGG 240  
 OY 351 CGCCGAGCTCACCGCCTTAACGTCGTGCGGAGCAATGSGGCTCTCAGAGAGATGCCGCGCAGCT 410  
 Db 241 CGCCGAGCTCACCGCCTTAACGTCGTGCGGAGCAATGSGGCTCTCAGAGAGATGCCGCGCAGCT 300  
 OY 411 GCAGGCGCGCAGCGACGACGAGG 431  
 Db 301 GCAGGCGCGCAGCGACGACGACGAGG 321  
 RESULT 6  
 B1756931/c 320 bp mRNA linear EST 25-SEP-2001  
 LOCUS 603029136F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5199417 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1756931  
 VERSION B1756931.1 GI:15748509  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 320)  
 NIH-MGC <http://mgc.ncl.nih.gov/>.  
 TITL E National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgsabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution by: MGC clone distribution Information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLML1499 row: h column: 10 High quality sequence stop: 320.
	FEATURES	Location/Qualifiers
	source	1..320
		/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5199417" /clone_lib="NIH_MGC_114" /lab_host="DH10B" /note="Organ: brain; Vector: pCMV-Sport6; Site:1; Notif: site:2; EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH-MGC library."
BASP COUNT	43 a 122 c 102 g 53 t	
ORIGIN		
Query Match	51.1%;	Score 320; DB 13; Length 320;
Best Local Similarity	100.0%;	Pred. No. 8; 6e-50;
Matches 320; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
OY 117	ACGGGAGTGACGGCGCAGCAAGCGCGGGGATCTCGAGCCATGAGGGCGCGCGCGACGC 176	
Db 320	ACGGGGGTGACGGCGCGCAGCGCGCGGGGATCTCGAGCCATGAGGGCGCGCGCGACGC 261	
OY 177	CATCTCGTAGTCGCTGTGAAGAACCCTACGCCGAGAGACTCAAGAAATTCAAAGCTGAAGCT 236	
Db 260	CATCTCGTAGTCCGTGTGAAGAACCCTACGCCGAGAGACTCAAGAAATTCAAAGCTGAAGCT 201	
OY 237	GCTGTGCGTGCGCGTGCAGAGAGGCTACGGGCGGCATCCGCGGGCGCGCTGTGTCCAT 296	
Db 200	GCTGTGCGTGCGCGTGCAGAGAGGCTACGGGCGGCATCCGCGGGCGCGCTGTGTCCAT 141	
OY 297	GGACGCTTGGACTCACCGAACAACCTGTGTAGCTTCTACGTGAGAACCTACGGCGCGCA 356	
Db 140	GGACGCTTGGACTCACCGAACAACCTGTGTAGCTTCTACGTGAGAACCTACGGCGCGCA 81	
OY 357	GCTCACCGCTTAACGTGTCGCGAGACATGGGCTTCCAGAGAGATGGCGGGCGACGTGCAGGC 416	
Db 80	GCTCACCGCTTAACGTGTCGCGAGACATGGGCTTCCAGAGAGATGGCGGGCGACGTGCAGGC 21	
OY 417	GGCACGCGACACAGGGTGAAGC 436	
Db 20	GGCACGCGACACAGGGTGAAGC 1	
RESULT 7		
BM459678		
LOCUS	BM459678	930 bp mRNA linear EST 05-FEB-2002
DEFINITION	AGNCNURP_6417789 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534308	
VERSION	BM459678	
KEYWORDS	BM459678.1 GI:18508718	
SOURCE	EST.	
ORGANISM	human.	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 930)	
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	



TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLM12769 row: c column: 22  
High quality sequence stop: 570.  
Location/Qualifiers  
1. .792  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="5745741"  
/clone\_lib="NIH\_MGC\_118"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source: Leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed for cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH\_MGC Library."

BASE COUNT 182 a 241 c 242 g 122 t 5 others

ORIGIN

Query Match 50.5%; Score 316; DB 13; Length 792;  
Best Local Similarity 100.0%; Pred. No. 4.2e-49;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 CACGGGGTGAACGGCGGACAGCGCCGGGAGTCTGAGCCATGGGGCGCCGGCGAGC 175  
|||||  
Db 26 CACGGGGTGAACGGCGGACAGCGCCGGGAGTCTGAGCCATGGGGCGCCGGCGAGC 85  
|||||  
QY 176 CCATCGTGGATGGGTGGAGAACTGACCGCGAGAGCTCAAGAACTCAAGTGAAGC 235  
|||||  
Db 86 CCATCGTGGATGGGTGGAGAACTGACCGCGAGAGCTCAAGAACTCAAGTGAAGC 145  
|||||  
QY 236 TGTGTGCGGTGCGGTGCGGAGGAGGCTACGCGCATCCGGGGCGCGTGTCTCA 295  
|||||  
Db 146 TGTGTGCGGTGCGGTGCGGAGGAGGCTACGCGCATCCGGGGCGCGTGTCTCA 205  
|||||  
QY 296 TGGACGCGTTGGACCTGACGACAAGCTGTGAGCTTACTGAGAGCTACAGCGCGCG 355  
|||||  
Db 206 TGGACGCGTTGGACCTGACGACAAGCTGTGAGCTTACTGAGAGCTACAGCGCGCG 265  
|||||  
QY 356 AGCTCACCGCTTAAGCTGCTGCGGAGATGGGCGCTGAGAGAGATGAGCGCGAGCTGCA 415  
|||||  
Db 266 AGCTCACCGCTTAAGCTGCTGCGGAGATGGGCGCTGAGAGAGATGAGCGCGAGCTGCA 325  
|||||  
QY 416 CGGCGACGACGAGG 431  
|||||  
Db 326 CGGCGACGACGAGG 341  
|||||

RESULT 10  
BM926438 804 bp mRNA linear EST 12-MAR-2002  
LOCUS BM926438  
DEFINITION AGENCOURT 6644759 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5766806  
5', mRNA sequence.  
ACCESSION BM926438  
VERSION BM926438.1 GI:19376817  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 804)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLM12824 row: a column: 15  
High quality sequence stop: 746.  
Location/Qualifiers  
1. .804  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="5766806"  
/clone\_lib="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source: anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

BASE COUNT 165 a 238 c 243 g 125 t 33 others

ORIGIN

Query Match 50.3%; Score 315; DB 14; Length 804;  
Best Local Similarity 100.0%; Pred. No. 6.5e-49;  
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 ACGGGGTGAGCGGCGGACGCGCGGGATCTGAGACCTGAGGGCGCGGAGAGC 176  
|||||  
Db 16 ACGGGGTGAGCGGCGGACGCGCGGGATCTGAGACCTGAGGGCGCGGAGAGC 75  
|||||  
QY 177 CATCTGGATGCGCTGAGAACTGACCGCGAGAGCTCAAGAGTTCAAGCTGAAGCT 236  
|||||  
Db 76 CATCTGGATGCGCTGAGAACTGACCGCGAGAGCTCAAGAGTTCAAGCTGAAGCT 135  
|||||  
QY 237 GCTGTGCGGTGCGGTGCGGAGGAGCTACGCGCATCCGGGGCGCGTGTCCAT 296  
|||||  
Db 136 GCTGTGCGGTGCGGTGCGGAGGAGCTACGCGCATCCGGGGCGCGTGTCCAT 195  
|||||  
QY 297 GGACGCGTTGACCTGACGACAAGCTGTGAGCTTACTGAGAGACTACGCGCGCA 356  
|||||  
Db 196 GGACGCGTTGACCTGACGACAAGCTGTGAGCTTACTGAGAGACTACGCGCGCA 255  
|||||  
QY 357 GCTCACGCGCTTAAGCTGCTGCGGAGATGGGCGCTGAGAGAGATGAGCGCGAGCTGCA 416  
|||||  
Db 256 GCTCACGCGCTTAAGCTGCTGCGGAGATGGGCGCTGAGAGAGATGAGCGCGAGCTGCA 315  
|||||  
QY 417 GGCGACGACGAGG 431  
|||||  
Db 316 GGCGACGACGAGG 330  
|||||

RESULT 11  
BC764161 750 bp mRNA linear EST 15-MAY-2001  
LOCUS BC764161  
DEFINITION 602737190F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4862182 5',  
mRNA sequence.  
ACCESSION BC764161  
VERSION BC764161.1 GI:14074814  
KEYWORDS EST.

ORGANISM	SOURCE
human.	human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	
1 (bases 1 to 750)	
NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .	
NIH-MGC	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	
Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a>	
Tissue Procurement: ATCC/DCMP/DRP	
cDNA Library Preparation: Ling Hong/Rubin Laboratory	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
DNA Sequencing by: Incyte Genomics, Inc.	
Clone distribution: MGC clone distribution information can be	
found through the I.M.A.G.E. Consortium/LLNL at:	
<a href="http://image.llnl.gov">http://image.llnl.gov</a>	
Plate: LCM1722 row: g column: 23	
High quality sequence stop: 721.	
FEATURES	
source	
1..750	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:4862182"	
/clone_lib="NIH_MGC_49"	
/tissue_type="melanotic melanoma, high MDR (cell line)"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: skin; Vector: pORF7; Site.1: XhoI; Site.2:	
EcoRI; cDNA made by oligo-dT priming. Directionally cloned	
into EcoRI/XhoI sites using the following 5' adaptor:	
GGCAGCAG(G). Size-selected >500bp for average insert size	
1.8kb. Library constructed by Ling Hong in the laboratory	
of Gerald M. Rubin (University of California, Berkeley)	
using ZAP-cDNA synthesis kit (Stratagene) and Superscript	
II RT (Life Technologies). Note: this is a NIH_MGC	
Library.  "	
BASE COUNT	
165 a 232 c 235 g 118 t	
ORIGIN	
Query Match	50.2%; Score 314; DB 12; Length 750;
Best Local Similarity 100.0%; Pred. No. 1e-48;	
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY 118 GCGGGGTGAGCGGGGCGAGCGCGGGGATCCTGAGACCTAGGGGGCGCGCGGAGAGCC 177	
DB 2 GCGGGGTGAGCGGGGCGAGCGCGGGGATCCTGAGACCTAGGGGGCGCGCGGAGAGCC 61	
OY 178 ATCTGATGATGCGCTGAGAACTGACCGCGGAGAGAGCTCAAGAAATTCAAGCTGAAGCTG 237	
DB 62 ATCTGATGATGCGCTGAGAACTGACCGCGGAGAGAGCTCAAGAAATTCAAGCTGAAGCTG 121	
OY 238 CTGTGCGTGGCGGTGCGGAGGCGCTACGGGCGCAATCCGGGGGGGCGCGGTGCTGCATG 297	
DB 122 CTGTGCGTGGCGGTGCGGAGGCGCTACGGGCGCAATCCGGGGGGGCGCGGTGCTGCATG 181	
OY 298 GAGCGCTTGAGACTTACCGCAAGCTGCTGACCTTCTTACTCTGAGAGACTTACCTGAGGCGGAG 357	
DB 182 GAGCGCTTGAGACTTACCGCAAGCTGCTGACCTTCTTACTCTGAGAGACTTACCTGAGGCGGAG 241	
OY 358 CTGACCGCTTAAAGCTGCTGCGGAGCATGGGCGTGGAGAGATGCGCGGAGCGTGGAGAGCG 417	
DB 242 CTGACCGCTTAAAGCTGCTGCGGAGCATGGGCGTGGAGAGATGCGCGGAGCGTGGAGAGCG 301	
OY 418 GCCACGCACACAGG 431	
DB 302 GCCACGCACACAGG 315	
RESULT 12	
BE67422	
LOCUS	BC67422
DEFINITION	6072412191 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4871020 5',
mRNA sequence.	856 bp mRNA linear EST 15-MAY-2000

	Accession	B6767422
	Version	B6767422.1 GI:14078075
	Keywords	EST.
	Source	human.
	Organism	Homo sapiens
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	Reference	1 (bases 1 to 856)
	Authors	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
	Title	National Institutes of Health, Mammalian Gene Collection (MGC)
	Journal	Unpublished (1999)
	Comment	Contact: Robert Strussberg, Ph.D. Email: gcaps-remail.nih.gov Tissue Procurement: ATCC/DCTP/DRP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution by: MGC clone distribution Information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LNCM1745 row: e column: 05 High quality sequence stop: 696. Location/Qualifiers
FEATURES	source	1..856
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:4871020"
		/clone_1lb="NIH-MGC_49"
		/tissue.type="melanotic melanoma, high MDR (cell line)"
		/lab_host="DH10B (phage-resistant)"
		/note="Organ: skin; Vector: pORF7; site_1: XhoI; site_2:
		EcoRI; cDNA made by oligo-dr priming. Directionally cloned
		into EcoRI/XhoI sites using the following 5' adaptor:
		GCGCAG(G). Size-selected >500bp for average insert size
		1.8kb. Library constructed by Ling Hong in the laboratory
		of Gerald M. Rubin (University of California, Berkeley)
		using ZAP-cDNA synthesis kit (Stratagene) and Superscript
		II RT (Life Technologies). Note: this is a NIH-MGC
		Library."
BASE COUNT	194 a 234 c 279 g 149 t	
ORIGIN		
Query Match	50.2%; Score 314; DB 12; Length 856;	
Best Local Similarity	100.0%; Pred. No. 9, 8e-49;	
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	118 GCGGGGTAGGCGCGCGAAGCGCGCGGGAATCTTGAGCAATGGGGGCGCGCGAGGCC	177
DB	2 GCGGGGTAGGCGCGCGAAGCGCGCGGGAATCTTGAGCAATGGGGGCGCGCGAGGCC	61
OY	178 ATCTGTGAATGCTGTGAAACCTGACCCTCCAGAGACTCAAGAAGTTCAAAGCTGAAGCTG	237
DB	62 ATCTGTGAATGCTGTGAAACCTGACCCTCCAGAGACTCAAGAAGTTCAAAGCTGAAGCTG	121
OY	238 CTTGTGGTGGTCGTCGCGGAGGGGCTACGGGGCAATCCCGGGGGCCGGCTGTGCACATG	297
DB	122 CTGTGGTGGTCGTCGCGGAGGGCTACGGGGCAATCCCGGGGGCCGGCTGTGCATG	181
OY	298 GAGCGCTTGAGACTCACGACAAGCTGTGACATTCTTAACCTGAAGACTTACGGCGCCGAG	357
DB	182 GAGCGCTTGAGACTCACGACAAGCTGTGACATTCTTAACCTGAAGACTTACGGCGCCGAG	241
OY	358 CTCACGGCTTAACCTTCGCGGACATGGGGCTCGAGAGATGGCGGGGCAAGCTGAGAGCG	417
DB	242 CTCACGGCTTAACCTTCGCGGACATGGGGCTTCGAGAGATGGCGGGGCAAGCTGAGAGCG	301
OY	418 GCCACGACACAGG 431	
DB	302 GCCACGACACAGG 315	

LOCUS BG337806 748 bp mRNA linear EST 27-FEB-2001  
DEFINITION 602435709F1 NIH\_MGC\_46 Homo sapiens cDNA clone IMAGE:4553378 5',  
mRNA sequence.  
ACCESSION BG337806  
VERSION BG337806.1 GI:13144342  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 748)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
AUTHORS  
TITLE  
COMMENT  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
plate: LLM1251 row: b column: 03  
High quality sequence stop: 710.  
Location/Qualifiers  
1. 748  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4553378"  
/clone\_1b="NIH\_MGC\_46"  
/tissue\_type="telomysarcoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pOT87; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using 2AP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

BASE COUNT 153 a 239 c 238 g 118 t  
ORIGIN  
Query Match 49.9%; Score 312.4; DB 12; Length 748;  
Best Local Similarity 99.7%; Pred. No. 2e-48;  
Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 118 GCGGGGTGACGGCGGCGGCGGATCTCTGAGCCATGGGCGCGCGCGACGCC 177  
|||||  
Db 2 GCGGGGTGACGGCGGCGGCGGATCTCTGAGCCATGGGCGCGCGCGACGCC 61  
QY 178 ATCTGTGATGCGGTGAGAACTGACCGCCGAGGAGCTCAAGAGTTCAAGCTGAAGCTG 237  
|||||  
Db 62 ATCTGTGATGCGGTGAGAACTGACCGCCGAGGAGCTCAAGAGTTCAAGCTGAAGCTG 121  
QY 238 CTGTGCGGTGCGGTGCGGAGAGGCTACAGGGGCGCATCCCGGGGGCGCGTGTCTCATG 297  
|||||  
Db 122 CTGTGCGGTGCGGTGCGGAGAGGCTACAGGGGCGCATCCCGGGGGCGCGTGTCTCATG 181  
QY 298 GAGCGCTTGGACCTCAACCAAGCTGCTAGCTTACTTACTGAGAGCTACGCGCGCGAG 357  
|||||  
Db 182 GAGCGCTTGGACCTCAACCAAGCTGCTAGCTTACTTACTGAGAGCTACGCGCGCGAG 241  
QY 358 CTCACCGCTTAAGTGTGTGCGGAGCATGGGCTGCAAGAGATGGCGGGAGCTGCAAGCG 417  
|||||  
Db 242 CTCACCGCTTAAGTGTGTGCGGAGCATGGGCTGCAAGAGATGGCGGGAGCTGCAAGCG 301  
QY 418 GCCACGACCAAGG 431  
|||||  
Db 302 GACACGACCAAGG 315

RESULT 14  
LOCUS BM918855 1294 bp mRNA linear EST 12-MAR-2002  
DEFINITION AGENCOURT 6635188 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5747943  
5', mRNA sequence.  
ACCESSION BM918855  
VERSION BM918855.1 GI:19369234  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1294)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
AUTHORS  
TITLE  
COMMENT  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
plate: LLM12774 row: c column: 16  
High quality sequence stop: 304.  
Location/Qualifiers  
1. 1294  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5747943"  
/clone\_1b="NIH\_MGC\_120"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCW-SORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for "full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 280 a 471 c 279 g 262 t  
ORIGIN  
Query Match 48.1%; Score 301.2; DB 14; Length 1294;  
Best Local Similarity 97.5%; Pred. No. 2.2e-46;  
Matches 306; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 118 GCGGGGTGAGCGGCGGCGGCGGATCTCTGAGCCATGGGCGCGCGCGACGCC 177  
|||||  
Db 12 GCGGGGTGAGCGGCGGCGGCGGATCTCTGAGCCATGGGCGCGCGCGACGCC 71  
QY 178 ATCTGTGATGCGGTGAGAACTGACCGCCGAGGAGCTCAAGAGTTCAAGCTGAAGCTG 237  
|||||  
Db 72 ATCTGTGATGCGGTGAGAACTGACCGCCGAGGAGCTCAAGAGTTCAAGCTGAAGCTG 131  
QY 238 CTGTGCGGTGCGGTGCGGAGAGGCTACAGGGGCGCATCCCGGGGGCGCGTGTCTCATG 297  
|||||  
Db 132 CTGTGCGGTGCGGTGCGGAGAGGCTACAGGGGCGCATCCCGGGGGCGCGTGTCTCATG 191  
QY 298 GAGCGCTTGGACCTCAACCAAGCTGCTAGCTTACTTACTGAGAGCTACGCGCGCGAG 357  
|||||  
Db 192 GAGCGCTTGGACCTCAACCAAGCTGCTAGCTTACTTACTGAGAGCTACGCGCGAG 251  
QY 358 CTCACCGCTTAAGTGTGTGCGGAGCATGGGCTGCAAGAGATGGCGGGAGCTGCAAGCG 417  
|||||  
Db 252 CTCACCGCTTAAGTGTGTGCGGAGCATGGGCTGCAAGAGATGGCGGGAGCTGCAAGCG 311  
QY 418 GCGACGACCAAGG 431  
|||||



Db 312 TCAAGCATATG 325

# RESULT 15 BM726665

DEFINITION BM726665 446 bp mRNA linear EST 01-MAR-2002  
UI-E-EJ0-ai1-n-17-0-UI.x1 UI-E-EJ0 Homo sapiens cDNA clone  
UI-E-EJ0-ai1-n-17-0-UI 5', mRNA sequence.

ACCESSION BM726665  
KEYWORDS BM726665.1 GI:19047998  
EST.

## SOURCE ORGANISM

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE AUTHORS

1 (bases 1 to 446)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene  
discovery

Genome Res. 6 (9), 791-806 (1996)

## JOURNAL MEDLINE COMMENT

Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

## FEATURES source

Location/Qualifiers  
1 446

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-E-EJ0-ai1-n-17-0-UI"

/clone\_11b="UI-E-EJ0"

/tissue\_type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
Choroid"

/dev\_stage="fetal and adult"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;

UI-E-EJ0 is a subtracted cDNA library constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT73-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tags for this library are: fetal eyes, AGAATCAGCA

; lens, CGATTAGCA; eye anterior segment, AATCCGCAT;

optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and

Macular, GTCC; RPE and Choroid, ACCCA. This library was

created for the program, Gene Discovery in the Visual

System, supported by National Eye Institute (NEI)."

BASE COUNT 76 a 145 c 156 g 69 t  
ORIGIN

## Query Match

47.3%; Score 296; DB 14; Length 446;

Best Local Similarity 100.0%; Pred. No. 2.3e-45;

Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 GCGGCGGGAATCCTGGAGCATGGGGCGGCGGCGAGCATCTGGATGCGCTGGAG 195

Db 1 GCGGCGGGAATCCTGGAGCATGGGGCGGCGGCGAGCATCTGGATGCGCTGGAG 60

Qy 196 AACCTGACCGCCGAGAGCTCAAGATTCAAGCTGCTGTGCGTGGCGGCG 255  
|||||  
Db 61 AACCTGACCGCCGAGAGCTCAAGATTCAAGCTGCTGTGCGTGGCGGCG 120  
Qy 256 GAGGGCTACGGGCGCATCCCGCGGGCGCGCTGCTGCATGAGACGCTTGACCTACC 315  
|||||  
Db 121 GAGGGCTACGGGCGCATCCCGCGGGCGCGCTGCTGCATGAGACGCTTGACCTACC 180  
Qy 316 GACAGCGTGTACGCTTCTACCTGAGACCTACGGCGCGGAGCTCAACGCTGCTG 375  
|||||  
Db 181 GACAGCGTGTGTAGCTTCTTCTACCTGAGACCTACGGCGCGGAGCTCAACGCTGCTG 240  
Qy 376 CGGCAATGGGCTTGCAGAGATGGCCGGGCAAGCTGCAGCGGCGCACCGACGAG 431  
|||||  
Db 241 CGGCAATGGGCTTGCAGAGATGGCCGGGCAAGCTGCAGCGGCGCACCGACGAG 296

Search completed: December 22, 2002, 06:45:56  
Job time : 1999 sees

